SITE-SPECIFIC STRUCTURE AND DYNAMICS OF POLYGLUTAMINE-CONTAINING AMYLOID FIBRILS AND THE CAVEOLIN SCAFFOLDING DOMAIN BY MAGIC ANGLE SPINNING SOLID-STATE NMR

by

Cody Lee Hoop

BA in Chemistry, Washington & Jefferson College, 2008

Submitted to the Graduate Faculty of

The School of Medicine in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2014

UNIVERSITY OF PITTSBURGH

SCHOOL OF MEDICINE

This dissertation was presented

by

Cody Lee Hoop

It was defended on

November 24, 2014

and approved by

Dr. W. Seth Horne, Assistant Professor, Department of Chemistry

Dr. Rieko Ishima, Associate Professor, Department of Structural Biology

Dr. Patrick H. Thibodeau, Assistant Professor, Department of Cell Biology

Dr. Ronald Wetzel, Professor, Department of Structural Biology

Dissertation Advisor: Dr. Patrick C.A. van der Wel, Assistant Professor, Department of Structural Biology

SITE-SPECIFIC STRUCTURE AND DYNAMICS OF POLYGLUTAMINE-CONTAINING AMYLOID FIBRILS AND THE CAVEOLIN SCAFFOLDING DOMAIN BY MAGIC ANGLE SPINNING SOLID-STATE NMR

Cody Lee Hoop, PhD

University of Pittsburgh, 2014

Protein assemblies and membrane associations play important roles in health and disease. Structural studies of these biological complexes are thus vital, but are often unfeasible by conventional tools in structural biology. Solid-state NMR (ssNMR) has made possible atomic-level structural studies of large, insoluble, and non-crystalline biological systems, such as amyloid fibrils and membrane proteins.

Amyloid fibrils are associated with at least 20 human diseases, making structural aspects of their formation crucial to understanding their aggregation pathways. Huntington's disease (HD) is caused by an expansion beyond a threshold of the polyglutamine (polyQ) domain in the Exon 1 of the huntingtin protein (htt). The 17-residue N-terminal segment of the Exon 1 (htt^{NT}) initiates the fibril aggregation and helps stabilize oligomers and fibrils. On the contrary, a polyPro segment C-terminal to the polyQ reduces fibril aggregation. In this thesis, magic-angle spinning (MAS) ssNMR was used to elucidate atomic-resolution structure and dynamics of the polyQ and its flanking domains in polyQ-containing amyloid-like fibrils.

More than 25% of human proteins are membrane proteins. Caveolin-1 (Cav1) is a protein found associated with cholesterol-rich membranes that forms caveolae, curved cave-like invaginations in the plasma membrane, and is implicated in muscular diseases and cancers. The caveolin-scaffolding domain (CSD) of Cav1 is responsible for cholesterol-recognition and

oligomerization to form the caveolae. In this thesis, MAS and static ssNMR are used to elucidate molecular structure of the CSD and its perturbation of the lipid bilayer in a cholesterol-rich lipid environment. These results allow for a more thorough understanding of the role of CSD in caveolae formation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

First, I would like to thank my dissertation advisor, Dr. Patrick van der Wel, for his mentorship, patience, and motivation over the past six years. I thank him for having faith in me from the very beginning, even before moving to Pittsburgh. From him, I have certainly learned a great deal about structural biology studies, especially in ssNMR, but how to do good science and scientific communication are among the greatest lessons he has passed on. I would also like to thank my first research advisor, Dr. Robbie Iuliucci, for getting me hooked on ssNMR as a naïve undergraduate. His inspiration brought me to the University of Pittsburgh in the first place. I am also grateful for his continued mentorship even after my undergraduate studies. I would also like to thank my committee for all of their helpful advice and insights throughout this dissertation.

My fellow MBSB colleagues have been great friends and teachers during the past few years. I would especially like to thank my core support system, Drs. Wazo Myint and Elizabeth (Landrum) Yu. They were always there at least to be cheerleaders through my work here (pom-poms and all). They were always available to discuss problems, no matter scientific or personal in nature. All of the current and former members of the van der Wel lab have been a tremendous source of scientific knowledge. I could always count on them for help with experimental methods and concepts. I also acknowledge the members of the Wetzel lab for their continuous support through collaborations, especially with sample preparation, taking EM images, and

measuring FTIR spectra. Their expertise with the fibril samples was a great asset to completing this thesis.

Last, but certainly not least, I would like to thank my close friends and family for their continued support and interest in my work, even if their only vision of my work day was through the eyes of the guys on The Big Bang Theory.

1.0 INTRODUCTION

1.1 MACROMOLECULAR ASSEMBLIES

In the study of macromolecular assemblies, solid-state nuclear magnetic resonance (ssNMR) has become an important tool to study structure and dynamics. The aim of this thesis is to elucidate structural and dynamical details at atomic-level resolution of the polyQ-amyloid core and its flanking segments in mature amyloid-like fibrils and the oligomerization domain, CSD, in the Cav1 membrane-associated protein. The insoluble and non-crystalline nature of amyloid fibrils makes them impossible to study by solution NMR or X-ray crystallography, making ssNMR the prime biophysical technique for their structural elucidation. The ability to study membrane proteins in context of lipid bilayers allows for more biologically relevant studies by ssNMR of these systems. In this section, an overview of the types of macromolecular assemblies studied in this thesis is presented.

1.1.1 Amyloid Fibrils

Amyloid fibrils are insoluble, filamentous, β-sheet rich peptide or protein aggregates. Misfolding and aggregation of proteins is associated with at least 20 human disorders, including Alzheimer's disease, Parkinson's disease, and Huntington's disease [1]. Understanding their

respective fibril aggregation pathways will aid in therapeutic design. Though disease toxicity seems correlated to protein misfolding and aggregation, the toxic species is unclear, and is often believed to be oligomeric intermediates [2, 3]. For a thorough understanding of the molecular causes of such diseases, it is necessary to gain insight into structural changes of amyloid-forming proteins along the fibril formation pathway. In this work, I have concentrated my efforts on the mature fibrils, for which resolving structural details can aid in understanding the aggregation pathway.

The most characteristic feature of amyloid fibrils is the common cross-β pattern, identifiable by X-ray fiber diffraction. In the cross-β structure, backbone hydrogen-bonding running parallel to the long fibril axis stabilizes β-sheets with an inter-strand spacing of ~4.8 Å [4]. The backbone hydrogen bonding along the fibril axis is evident in **Figure 1.1**. The organization of such β-sheets varies with protein. The most common arrangement is in-register parallel, in which residues form intermolecular hydrogen bonds to the same residue in the neighboring strand. This is the fibrillar assembly observed, for instance, in the immunoglobulin binding domain B1 of streptococcal protein G (GB1) [5] and is shown in the 17-42 peptide fragment of amyloid β [6] in **Figure 1.1**.

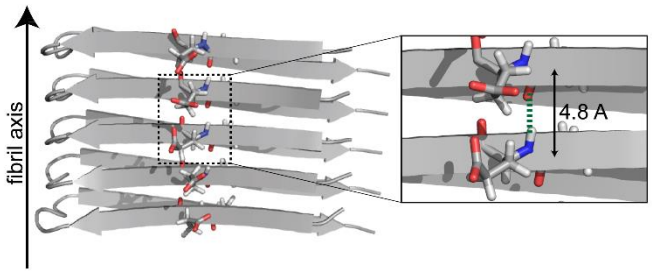

Figure 1.1 The cross-β structure of amyloid fibrils.

The 17-42 peptide fragment of amyloid β is used as an example (PDB: 2BEG) [6]. The black arrow on the left designates the fibril long axis. Backbone hydrogen bonds (green) run parallel to the fibril axis. The gray arrows

represent β-strands running from N-terminal to C-terminal ends. The in-register parallel assembly is evident from the alignment of the same residues along the fibril axis (in-register) as the stacked β-strands proceed in the same directions (parallel). The zoomed-in region shows the backbone hydrogen bonding between β–strands separated by ~4.8A.

These sheets often contain β-strands connected through short loops and stabilized through intramolecular side chain interactions. One of the most recurring strand-turn-strand motifs in amyloid proteins is the β-arcade, e.g. in amyloid β (Aβ) [6] (**Figure 1.1**) and HET-s [7] fibrils. Backbone hydrogen bonding in β-arcades occurs intermolecularly. Although parallel arrangements are more common in amyloids, β-arcades can also be arranged into antiparallel β-sheets (**Figure 1.2b**). Alternative to the arcade motif, a β-hairpin is another possible turn-containing structure, which is in the more common strand-turn-strand motif observed in globular protein structures (**Figure 1.2a**). Here, β-strands are stabilized through intramolecular backbone hydrogen bonds parallel to the fibril axis, and the turn between the β-strands often takes form of a β-turn.

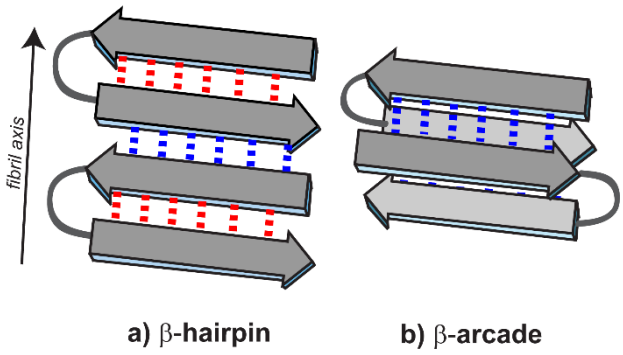

Figure 1.2 Turn structures in amyloid fibrils in an antiparallel arrangement.

a) A β-hairpin is a strand-turn-strand motif in which β-strands are stabilized by both intra- (red) and inter-molecular (blue) backbone hydrogen bonding along the fibril axis. b) A β-arcade is a strand-turn-strand motif that relies on intermolecular (blue) backbone hydrogen bonding.

It is common for amyloid-forming peptides to assemble into fibrils with different organizations, or polymorphs [8-11]. Polymorphs can form from the assembly of different numbers of filaments to form the fibril [12] or variations in their molecular structure, such as parallel or anti-parallel arrangement of β-sheets (as in the Iowa mutant of Aβ) [13] or the arrangement of backbone stabilizing hydrogen bonds [14]. Individual polymorphs can sometimes be enhanced through seeding of a particular morphology. By adding a small amount of mature fibrils, or seeds, to a solution of the same peptide or protein, the morphology of the seeds is carried on into the mature fibrils. Through several generations of seeding, the mature fibrils will consist of only a single morphology [15]. Polymorphism may be related to the pathology of disease. Tycko and colleagues have seeded fibril formation of $A\beta_{40}$ with sonicated brain extracts from different parts of the brain of different Alzheimer's disease patients to probe the structural features of the resultant fibrils [16]. They found that extracts from different parts of the same brain yielded fibrils with a common morphology but that the extracts from different patients seeded different structural polymorphs from each other. Research in imaging is ongoing to design probes to recognize specific amyloid conformations [17, 18].

Several studies have been investigating structural features of polyQ fibrils, including X-ray diffraction [19, 20], solid-state NMR [21, 22], and EPR [23]. It has been shown that polyQ fibrils tend to form in a less common arrangement of anti-parallel β-sheets [19, 20, 24]. Antiparallel β-hairpins, as in **Figure 1.2a**, have been suggested to play important roles in the aggregation of polyQ fibrils. It has been shown that polyQ fibril formation tolerates

incorporation of multiple PG turn enhancing segments within Gln stretches, provided there are a number of Gln residues between, without slowing aggregation [24]. Furthermore, incorporation of β-hairpin enhancers in polyQ fibrils actually increases aggregation rates and fibril stability [25]. By ssNMR, I have showed that the structure of these modified polyQ fibrils is similar to simple polyQ [25] (See **Section 1.3**). Yet, whether β-hairpins are present in the mature polyQ fibrils remains a topic of debate in the field. MAS ssNMR studies of residues in and around the turn in polyQ fibrils featuring β-hairpin enhancing Pro-Gly motifs are presented in **Chapter 3.0** .

In polyQ, the Gln residues interact through both backbones and side chains to stabilize the amyloid core. Several models have been proposed for the geometry of the Gln residues based on X-ray fibril diffraction. Perutz et al had concluded from X-ray fibril diffraction that $D_2Q_{15}K_2$ fibrils form a water-filled nanotube [26]. The Gln residues in this model all had the same side chain conformations with a χ_2 near 180^o. Based on a reinterpretation of this data, Sikorski and Atkins proposed that polyQ fibrils feature stacked β-strands connected by intramolecular β-hairpins, stabilized through hydrogen bonding and interdigitating extended Gln side chains in a steric zipper [19]. Another model shows the Gln side chains to be bent with a nearly 90^o χ_2 angle [20]. There has been a lack of detailed structural data required to resolve these inconsistencies between the models. In this thesis, extensive investigations of the structure and dynamics of the Gln residues in the amyloid core, including a new model of Gln conformers, are presented in **Chapters 2.0 and 3.0** .

1.1.2 Membrane-Associated Proteins: Caveolins

A membrane-associated protein found in cholesterol-rich environments, caveolin-1 (Cav1), is involved in muscular diseases and cancers [27]. Three caveolin proteins aid in the construction of caveolae: flask-shaped, cholesterol-rich invaginations in cell membranes that are involved in cell signaling, cell adhesion, and molecular transport [28, 29]. The caveolin proteins are found in caveolae-rich smooth-muscle cells, fibroblasts, endothelial cells, and adipocytes [29].

Formation of caveolae requires tight binding of caveolin to cholesterol [30]. SSNMR is a powerful tool to study proteins in an array of environments, including caveolin's native cholesterol-rich lipid bilayer, and is capable of probing both lipid and protein entities to site-specific resolution. The three caveolins have a characteristic intramembrane α-helical hairpin domain that does not traverse through the membrane such that both N- and C-termini are cytoplasmic [29]. Cav1 and caveolin-3 form caveolae through oligomerization and association with cholesterol-rich lipid domains that occurs at the caveolin scaffolding domain (CSD) [29, 31]. The tight binding of Cav to cholesterol is thought to occur at a putative cholesterol recognition/interaction amino acid consensus (CRAC) motif [28]. CRAC motifs have the sequence: L/V-X_{1-5}-Y-X_{1-5}-R/K, which is V_{94}TKYWFYR$_{101}$ [32] in the CSD of Cav1. Also within the CSD are residues required for protein binding, F_{92}TVT$_{94}$. These residues are included in the caveolin-binding motif (CBM)- like sequence of the CSD, which is rich in aromatic amino acids [33]. These motifs are normally found in protein binding partners of Cav, but the similarity in sequence within Cav brings to light the role of this motif in Cav homo-oligomerization. Despite the functional importance of this 20- amino acid-long domain, structural information on

this domain remains elusive, in particular in context of cholesterol-rich membranes. In **Chapter 4.0** , our structural studies of the CSD of Cav1 in cholesterol-rich lipid bilayers are discussed.

1.2 SOLID-STATE NMR EXPERIMENTS USED IN THESE STUDIES

In the study of amyloid fibrils and membrane proteins, ssNMR fills a void in structural biology not accessible by other methods. Solution NMR requires the sample be soluble at concentrations high enough for detectable data and becomes challenging for molecules or complexes larger than ~30 kDa in size. This size limitation is of no issue for X-ray crystallography, but the difficulty there is in obtaining high quality crystals. Such limitations are not a problem for ssNMR, making it a prime tool to study large, complex biological systems. These advantages of ssNMR have made it an important technique in obtaining information at atomic resolution of complex biomolecules, such as amyloid fibrils and membrane proteins. In this thesis, a gamut of ssNMR experiments were performed to answer important biological questions pertaining to amyloid fibril structure and assembly and membrane protein structure. These experiments were used to determine secondary structure, constrain distances between nuclei, constrain backbone and side chain torsion angles, measure site-specific dynamics, and observe lipid conformations.

Biological NMR is a multi-stage process that requires sample preparation, acquisition of various data sets, and data analysis to obtain structural information. **Figure 1.3** is a flowchart of the sequential stages of the biological ssNMR process. One must first prepare a sample that is labeled with NMR-visible isotopes (those with non-zero spins). The spin-1/2 nuclei are the most commonly studied by NMR. In proteins, ^{1}H, ^{13}C, and ^{15}N nuclei are observed. ^{13}C and ^{15}N have low sensitivity, being only 1.1% and 0.37% abundant in nature, respectively. To overcome this, samples are isotopically labeled with these isotopes either uniformly (all carbon and nitrogen

sites), selectively (in only select carbon and nitrogen sites), or site-specifically (only in certain residues). Uniform and selective labeling are be done during the expression of a protein. Site-specific labeling is done through incorporation of ^{13}C and ^{15}N-encriched amino acids into the proper residue in the sequence during peptide synthesis. In ssNMR, once the sample is in the state of interest, e.g. as mature fibrils (as in **Figure 1.3a**), incorporated into lipids, etc., it is packed into the sample holder, called the "rotor" (**Figure 1.3b**). The cap that seals the rotor has fins that allow it to spin inside the NMR probe. In this thesis, most samples were packed into rotors with an outer diameter of 3.2 mm (as shown in **Figure 1.3b**), which can spin up to 24 kHz. We then start obtaining MAS NMR spectra that have hopefully well-resolved peaks (**Figure 1.3c,d**). The position of a peak in the spectrum is defined by the resonance frequency of the nucleus, which is called the "chemical shift" and is reported in ppm. The first goal is to figure out which peaks are representative of which atoms and/or residues, and thus tabulate the chemical shifts of the isotopically labeled atoms (chemical shift assignments). In this thesis, ^{13}C and ^{15}N chemical shifts are recorded. **Section 1.2.2** discusses how we use chemical shifts to determine secondary structure in proteins and peptides. Once peaks can be attributed to specific atoms, further measurements are done to obtain structural constraints (e.g. distance measurements as in **Figure 1.3f**, torsion angle measurements, etc.) and dynamics measurements (e.g. relaxation measurements). The ways by which these structural constraints and dynamics measurements are obtained by NMR are described in **Sections 1.2.3 and 1.2.4**, respectively. Analyzing and piecing together these results allows us to answer biological questions, such as features of a molecular structure of the studied sample (representative cartoon in **Figure 1.3g**).

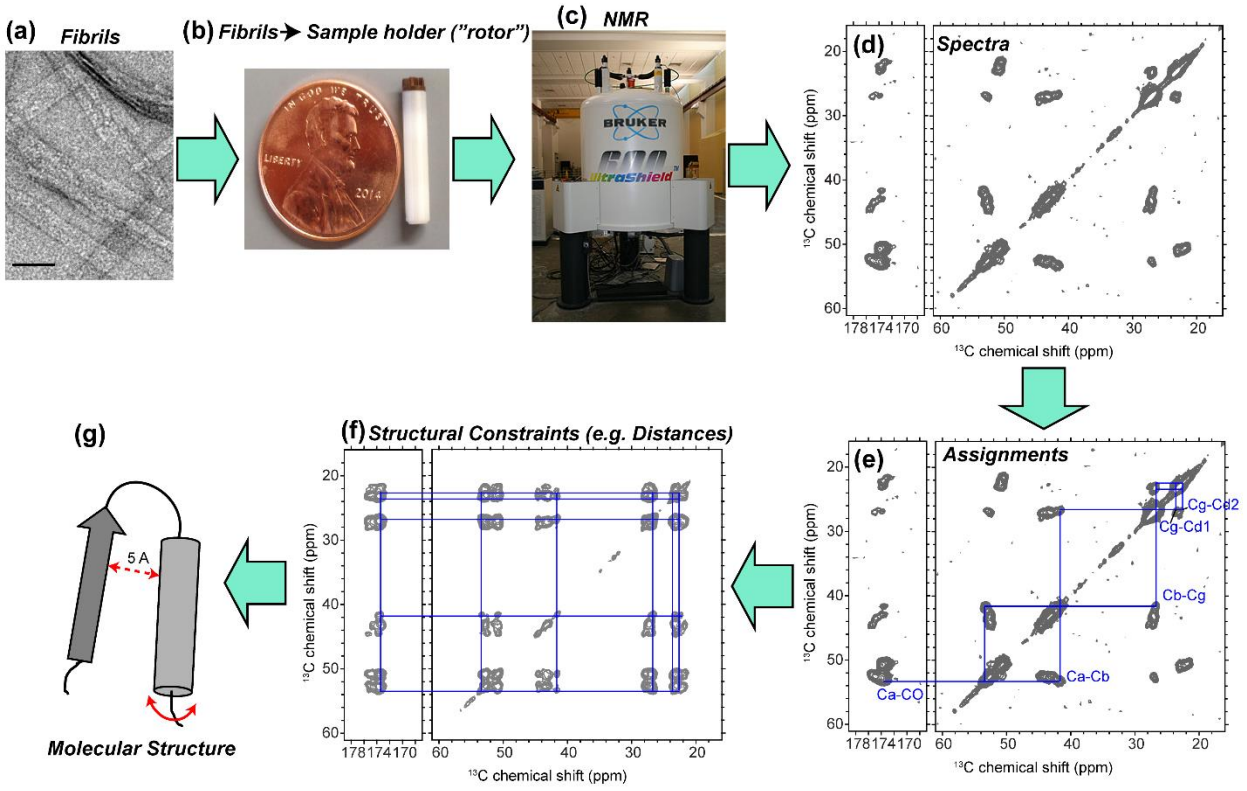

Figure 1.3 Basic workflow of biological ssNMR.

(a) Obtain a sample, shown here as fibrils. (b) Pack the sample into a rotor. The 3.2mm (outer diameter) shown here can contain at maximum 30 μL of sample and can spin up to 24 kHz in an MAS probe. (c) Place sample in the NMR magnet, and (d) acquire spectra. (e) Chemical shifts are assigned by correlating nearby ^{13}C atoms. (f) Further data is acquired to obtain structural constraints such as long distance contacts shown here. Piecing together these results provides answers to biologically interesting questions, such as molecular structure (g).

1.2.1 Obtaining Chemical Shift Assignments

An advantage to ssNMR is that structural and dynamic information can be obtained at the atomic level. Each atom (or more precisely, its nucleus) in a molecule gives a peak at a particular chemical shift. In two- and three-dimensional spectra, cross peaks intersecting at two chemical

shifts show an interaction between the two atoms to which those chemical shifts are assigned. Knowing the proximity of atoms to each other allows us to identify amino acids, map out connected residues, and even see intra- and inter-molecular contacts.

Rapidly spinning the sample during experiments at 54.7° to the external magnetic field, the "magic angle," averages the anisotropy from the slowly tumbling molecules and gives narrow peaks at the isotropic chemical shifts, which are comparable to those in solution NMR [34]. To obtain these chemical shifts, we use two- and three-dimensional ^{13}C-^{13}C and ^{15}N-^{13}C experiments, which rely on a fundamental building block of ssNMR, cross polarization (CP). CP is primarily used to enhance the sensitivity of low abundance nuclei (e.g. ^{13}C and ^{15}N) that are dipolar coupled to high abundance protons.

The technique primarily used for intra-residue ^{13}C-^{13}C correlations is DARR (Dipolar-Assisted Rotational Resonance) [35]. Proximate carbons will be observed at cross peaks at their chemical shifts. Upon extending the ^{13}C-^{13}C DARR mixing time used, carbons within ever longer distances will show a correlation. Often, an experiment with a short DARR mixing time, ~10 ms, is acquired early on to correlate carbons at most ~2.5 Å apart that are likely bonded and within the same residue in order to map out intraresidue interactions and assign those carbons to a specific amino acid type. Subsequent experiments with longer DARR mixing times, up to ~100 ms, will be acquired to observe longer ^{13}C-^{13}C distances. The red lines in the spectrum in **Figure 1.4**, which has intermediate DARR mixing time, shows all of the correlations within the Val residue in the N-Ac-VL peptide.

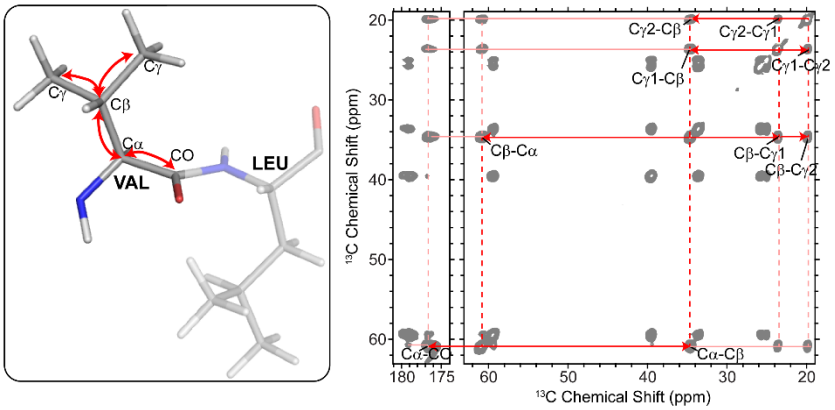

Figure 1.4 ^{13}C-^{13}C DARR used for intra-residue contacts.

The 2D ^{13}C-^{13}C DARR experiment is firstly used to connect carbons within residues. The red arrows on the dipeptide, VL, model show the magnetization transfers that occur with a short DARR mixing period. In the 2D spectrum on the right, those transfers are shown with solid red arrows, and the resulting peaks are labeled with their respective correlations. With increasingly long DARR mixing, e.g. 25ms as shown here, magnetization is transferred between further carbons. These are shown by pink lines in the 2D spectrum. Other peaks in the spectrum are attributed to Leu and an N-terminal acetyl group in the N-Ac-VL peptide.

After ^{1}H-^{13}C CP, an RF irradiation is applied to protons at a frequency equal to the spinning rate to enhance ^{13}C-^{13}C polarization transfer. **Figure 1.4** shows the magnetization transfer pathways for ^{13}C nuclei in the Val of dipeptide N-Ac-VL. The transfer of magnetization is manifest as a cross peak in a ^{13}C-^{13}C 2D spectrum.

In a protein or peptide, amino acids are connected by amide bonds between the carbonyl of one residue and the amide nitrogen of the next. ^{13}C chemical shift correlations are informative of the amino acid types, but not necessarily the specific residue within the primary sequence. We obtain and assign ^{15}N chemical shifts to connect sequential residues. The spectra in **Figure 1.5** show the ^{15}N-^{13}C correlations used to make ^{15}N assignments.

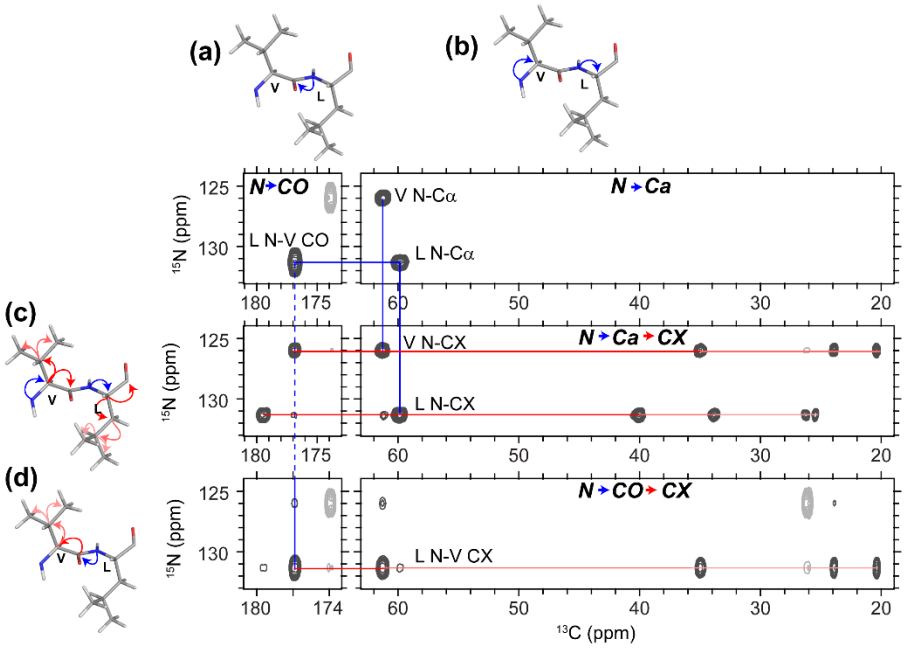

Figure 1.5 ^{15}N-^{13}C 2D and 3D experiments are used to connect adjacent residues.

The dipeptide VL is used here to show magnetization transfer pathways in ^{15}N-^{13}C assignment experiments. ^{15}N-^{13}C transfers are shown in blue and ^{13}C-^{13}C transfers in red. (a) NCO: in VL, the LN is correlated to the VCO. (b) NCA transfers magnetization from the amide nitrogen to its own Ca. Cross peaks are observed from VN-Cα and LN-Cα. (c) In NCACX, nitrogen is first correlated to the attached Cα, then to further carbons within the same residue. (d) In NCOCX, nitrogen is first correlated to the attached CO, then to carbons of the previous residue. Light gray cross peaks are from an N-terminal acetyl in the N-Ac-VL peptide.

Two- and three-dimensional ^{15}N-^{13}C correlation experiments are used to assign backbone and side chain nitrogens. In NCA and NCO experiments, ^{1}H-^{15}N CP first enhances the sensitivity of the ^{15}N nuclei. ^{15}N-^{13}C SPECIFIC CP (spectrally induced filtering in combination with cross polarization) [36] allows the selective transfer from ^{15}N to ^{13}Cα or ^{13}CO. The incorporation of these two CP steps is called double CP (DCP). The chemical shift is evolved on the ^{15}N and the selected ^{13}C nuclei are detected. For backbone assignments, NCA gives intra-residue amide nitrogen to Cα correlations, while NCO transfers the magnetization from the amide nitrogen to

the CO of the previous residue, as shown by the blue arrows in **Figure 1.5**. A subsequent ^{13}C-^{13}C DARR mixing period in NCACX and NCOCX experiments is used to transfer magnetization to nearby ^{13}C nuclei (red arrows in **Figure 1.5c and d**). ^{13}C-^{13}C DARR mixing periods of 10-50 ms allow for identification of residue types through the side chain chemical shifts. These residues are then connected together through the linking amide nitrogens. In a 3D NCACX or NCOCX, chemical shift is also evolved on Cα or CO, respectively, and is detected on ^{13}C nuclei.

In some cases, complementary one-, two-, or three-dimensional CANCO and/or CONCA correlations are also useful for identifying sequential residues (**Figure 1.6**).

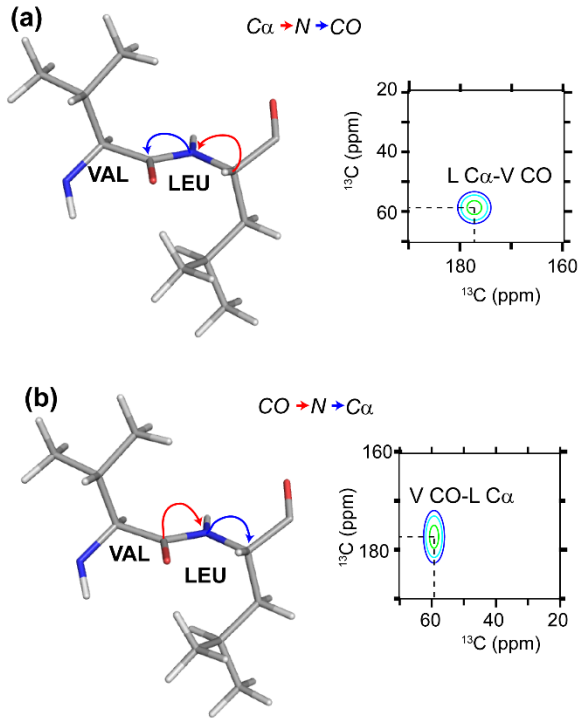

Figure 1.6 CANCO and CONCA connect adjacent residues.

Dipeptide VL is used to show magnetization transfer pathways of (a) CANCO and (b) CONCA experiments, shown here as 2D spectra. (a) CANCO first correlates Cα to its own amide nitrogen, and then to the attached CO of the preceding residue. A cross peak would be observed for LCα-VCO, as shown in the synthetic spectrum. (b) CONCA

first correlates CO to the subsequent amide nitrogen, and then to the attached CO of the subsequent residue. Here, a cross peak would be observed for VCO-LCα.

In these experiments, magnetization is first transferred from ^{1}H-^{13}C, instead of to ^{15}N. Magnetization is then selectively transferred from $^{13}C\alpha$ or ^{13}CO to ^{15}N via SPECIFIC CP. A second specific CP step is used to transfer magnetization from ^{15}N to ^{13}CO or $^{13}C\alpha$, respectively. The magnetization pathways are shown in **Figure 1.6**. Chemical shift is detected on the ^{13}C nuclei. As magnetization must be transferred through three CP steps, the sensitivity of these measurements is often low.

1.2.2 Determining Secondary Structure

For all 20 naturally occurring amino acids, characteristic patterns in chemical shifts of nuclei within and near the backbone have been observed to be dependent on the secondary structure. The method routinely used by biological NMR spectroscopists to evaluate the secondary structure from chemical shift is called Chemical Shift Indexing (CSI) [37]. Chemical shifts of all 20 amino acids in a random coil configuration, have been tabulated. The CSI method subtracts the random coil chemical shift from the observed chemical shift for a particular atom in a particular residue. This difference is referred to as the secondary chemical shift ($\Delta\delta$). $C\alpha$ and CO nuclei tend to be have a higher chemical shift relative to those of a random coil (positive $\Delta\delta$) in an α-helical structure and $H\alpha$ and $C\beta$ a lower chemical shift (negative $\Delta\delta$) [37]. The opposite occurs for residues in β-strand configurations. This method has been reported to be 75-85% in agreement with X-ray determined secondary structure [38]. Additionally, chemical shifts can be used to constrain backbone ϕ and ψ torsion angles through the use of TALOS (Torsion Angle

Likeliness Obtained from Shift and Sequence Similarity) [39]. This program compares the input chemical shifts to a database of chemical shifts of known structures and gives the backbone torsion angles for the closest matches of chemical shift and local primary sequence [39]. These backbone torsion angles are often plotted on a Ramachandran plot to help determine the secondary structure.

1.2.3 Obtaining Structural Constraints

Beyond secondary structure, inter-residue and inter-molecular interactions are key pieces to building a structure of a biomolecule. MAS ssNMR provides information on atomic proximities, which are indicative of potentially important intra- and inter-molecular interactions in a complex assembly. Typically, resonances can be observed between nuclei within 6 Å via Proton-Driven Spin Diffusion (PDSD).

During ^{13}C-^{13}C mixing, the proton decoupling is turned off allowing magnetization to transfer between ^{13}C nuclei via dipolar coupling. With long enough mixing times, up to ~500 ms, ^{13}C nuclei within ~6 Å are correlated by cross peaks. Buildup curves of the signal intensity over a range of ^{13}C-^{13}C mixing times are used to measure ^{13}C-^{13}C distances.

MAS ssNMR also provides us with tools to measure local geometries, such as torsion angles along the backbone and within side chains. With atomic distances and backbone and side chain torsion angles as constraints, a picture of the structure of a biomolecule, the architecture of a fibril assembly, the orientation of a protein in the membrane, etc. starts to emerge. The NCCN experiment derives the absolute value of the ψ backbone torsion angle. This angle is defined by the bonds between backbone atoms, N_i-$C\alpha_i$-CO_i-N_{i+1}. The pulse sequence is shown in **Figure 1.7a**.

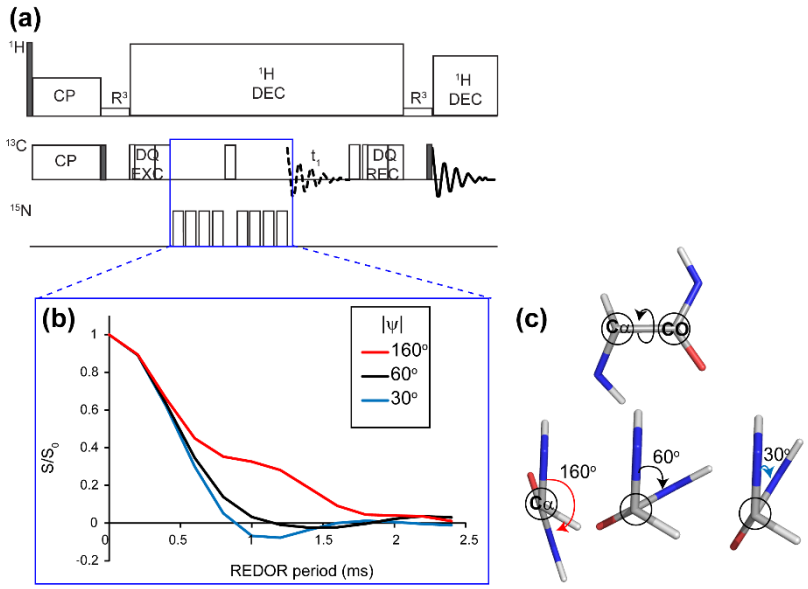

Figure 1.7 The NCCN experiment constrains the ψ backbone torsion angle.

(a) Pulse sequence schematic of the NCCN experiment. (b) Simulations of the dephasing of S/S_0 with an increasing REDOR ^{15}N recoupling period is shown for $|\psi|$ angles specified. The variable length REDOR period is shown in the blue box in (a). Simulations were conducted in the SPINEVOLUTION program [40]. (c) Schematics of the ψ dihedral angles 160°, 60°, and 30°, for which the simulated NCCN dephasing curves are shown in (b). The colors of the arrows denoting the angle are coordinated with the curves in (b).

The experiment creates a double quantum coherence between the bonded $^{13}C\alpha_i$ and $^{13}CO_i$ nuclei. This coherence is dephased due to ^{15}N-$^{13}C\alpha$ and ^{13}CO-^{15}N dipolar couplings. The dephasing curve is dependent on the relative orientation of the two N-C dipolar vectors to each other, as shown in **Figure 1.7b**, and thus the ψ torsion angle. In the NCCN experiment, a REDOR (Rotational Echo Double Resonance) period is used to recouple ^{15}N to ^{13}C with a series of rotor-synchronized π pulses on ^{15}N [41] (blue box in **Figure 1.7a**). The REDOR experiment measures dipolar coupling between heteronuclear spins, often ^{15}N-^{13}C [42]. Each data point along the dephasing curve is resultant from two sets of data: the experiment acquired with REDOR pulses

turned off for the indicated period length (S_0) and the analogous experiment with REDOR pulses turned on (S). The points on the dephasing curve are recorded as the ^{13}C signal intensity with REDOR recoupling normalized to the ^{13}C signal intensity without (S/S_0).

Through a similar experiment, side chain torsion angles are measured by dephasing the ^{13}C double quantum coherence through interference with dipolar-coupled protons, referred to as HCCH [43]. In this experiment, after excitation of a DQ coherence between bound ^{13}C nuclei, proton homonuclear decoupling is applied (e.g. Lee-Goldburg (LG) decoupling as in **Figure 1.8**) for a variable period of time between zero and one whole rotor period. The ^{13}C DQ intensity is then influenced by the H-C dipolar interactions. The interference pattern is indicative of the torsion angles that define the relative H-C bond orientations. **Figure 1.8c** is a schematic of the χ_2 torsion angle with only Cα, Cβ, Cγ, and Cδ and protons bound to the central carbons shown for simplicity. **Figure 1.8b** shows that the interference pattern from ^{1}H-^{13}C recoupling (blue arrow in **Figure 1.8a**) is dependent on the χ_2 torsion angle (Cα-Cβ-Cγ-Cδ).

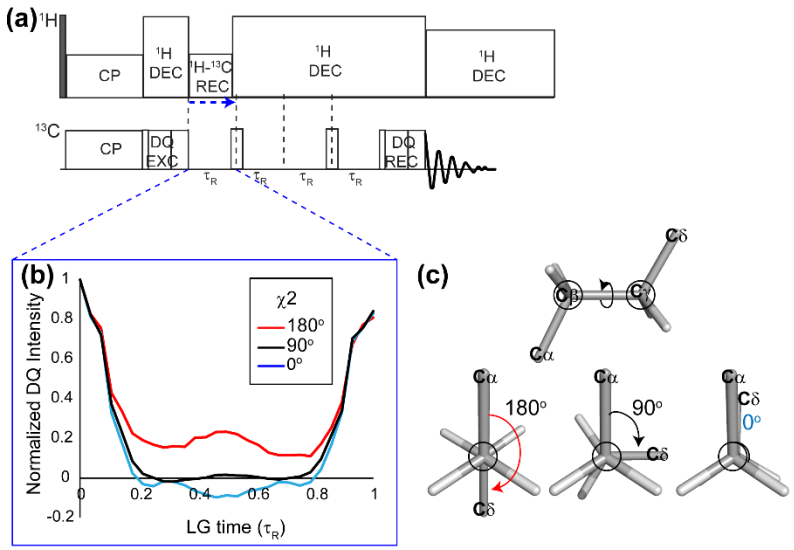

Figure 1.8 The HCCH experiment constrains dihedral angles around ^{13}C nuclei with attached protons.

(a) Pulse sequence schematic of the HCCH experiment. (b) Simulated interference patterns of the ^{13}C DQ intensity over an increasing period of LG homonuclear decoupling, measured in fractions of the rotor period (τ_R), for

indicated angles. Simulations were conducted in the SPINEVOLUTION program [40]. (c) χ2 torsion angles in a side chain, showing Cα, Cβ, Hβ1, Hβ2, Cγ, Hγ1, Hγ2, and Cδ. The HCCH experiment measures the orientation of the H-C dipoles to each other, which gives the χ2 angle, Cα-Cβ-Cγ-Cδ (arrows). The arrow colors correspond to their respective curves in (b).

Simulations of the pulse sequence under the employed experimental conditions are crucial for quantitative analysis of experiments such as those for measuring torsion angles, as above, or quantitating dynamics, discussed in **Section 1.2.4**. Simulations for quantitative analysis in this thesis were performed using the SPINEVOLUTION program [40]. In this program, the pulse sequence is input as pulse lengths and pulse power levels on appropriate channels as would be used in the experimental pulse sequence. A coordinate file for the system of interest is provided to the program, and each atom is specified as a type of nucleus (^{13}C, ^{15}N, ^{1}H). The user specifies the origin of magnetization and the spin to be observed. Selection of variables is dependent upon the quantitative information desired for a particular experiment. SPINEVOLUTION has also been used in this thesis for comparing the experimental data to the simulation results.

1.2.4 Measuring Site-Specific Dynamics

Dynamics play an integral role in protein stability, complex assembly, and biomolecular interactions. A qualitative measure of dynamics by ssNMR is the comparison of 1D ^{13}C experiments obtained through ^{13}C direct polarization (DP) versus dipolar-based ^{1}H-^{13}C CP, under otherwise identical conditions. This method has been used to gain insight into local dynamics in biological systems as the two polarization techniques have different sensitivities to dynamics.

The dipolar mediated ^{1}H-^{13}C CP transfer is less effective for mobile sites as the dipolar interaction is attenuated by dynamics. The DP experiment is dependent on ^{13}C relaxation and suppresses highly rigid sites due to their slow relaxation. Therefore, for rigid sites, the CP signal will dominate, while for mobile sites, the DP signal will dominate.

A more quantitative assessment of dynamics at atomic-level resolution in ^{13}C and ^{15}N nuclei is the measurement of longitudinal relaxation rates. Fast ps-ns motions are characterized by longitudinal relaxation rates (R_1). R_1 relaxation is the rate at which the z component of magnetization recovers to thermal equilibrium with its environment. The R_1 of ^{15}N or ^{13}C is measured as an exponential decay of signal intensity of a series of NMR experiments incorporating a variable length R_1 relaxation period. Both very mobile and rigid sites will have a low R_1, whereas intermediate mobility will cause a high R_1, measured in s^{-1}.

Another indicator of dynamics are amide backbone N-H dipolar couplings. The apparent N-H dipolar coupling strength is influenced by dynamics of the probed ^{15}N site. A high dipolar coupling constant is indicative of strong N-H dipolar coupling, and low dynamics, whereas a very mobile site is characterized by a weak N-H dipolar coupling and a low dipolar coupling constant. These can also be expressed as order parameters, where 1 is completely rigid and 0 indicates free motion. In **Chapter 2.0** , I have used rotor-synchronized R-symmetry sequences that recouple heteronuclear dipole-dipole interactions to measure N-H dipolar couplings in the polyQ amyloid core and the α-helix of the htt^{NT} and have estimated their respective order parameters.

1.2.5 Spectroscopic Filtering

Knowing the exposure of individual residues to solvent gives a sense of the compactness and surface area of the protein in a given state. Solvent exposure measurements give information on which residues are exposed to water at the surface and which are sequestered away from the solvent in compact structures. Typically, more mobile or less stable domains are more solvent exposed. We use relaxation filtering experiments in ssNMR to observe site-specific proximities to water. Mobile water protons have a long transverse relaxation time (T_2) relative to rigid protein protons. By inserting a T_2 filter long enough to eliminate protein protons and only allow through water protons (~3 ms) before the ^{1}H-^{13}C CP, ^{13}C nuclei can only be polarized ^{1}H polarization that originates from the water solvent. The intensity of magnetization transfer from water protons to ^{13}C nuclei is then indicative of the distance between the observed ^{13}C and water.

1.2.6 Analyzing Conformations of Lipids in Membrane Studies

MAS and static (without spinning) ssNMR have been particularly useful to study membrane-associated proteins in lipid environments. For most other applications used in this work, samples are spun at the magic angle, 54.7° relative to the external magnetic field, in order to eliminate chemical shift anisotropy (CSA) and obtain narrow isotropic peaks. However, underlying the CSA is important information relating to orientations and motions. Chemical shift anisotropy arises from the anisotropic electron distribution around a nucleus, thus giving a different chemical shift for a given nucleus depending on the molecular orientation. When molecules are randomly oriented with respect to the magnetic field, all possible chemical shifts of each nucleus in all orientations will be superimposed, giving a broad powder pattern. Static

ssNMR, without magic angle spinning allows information on orientation and motion to be extracted from this powder pattern. In static ssNMR, the ^{31}P line shapes are indicative of different lipid phases based on the relative orientations of phospholipid head-groups. As the phospholipids in a lipid bilayer are continuously in a state of anisotropic motion, a static 1D ^{31}P NMR will typically reveal a broad line shape representative of the motionally averaged CSA of the head-group. By analyzing this line shape, one can characterize the orientations and motions of the lipids and how they are affected by the presence of protein and in differing environments. ^{31}P NMR is used in the studies of the CSD in Cav1 in **Chapter 4.0** .

1.2.7 Summary of ssNMR Techniques

In summary, ssNMR is quite powerful in obtaining site-specific structural and dynamic information on biological complexes in a wide range of environments. Knowing which chemical shifts belong to which atoms and residues is already telling of secondary structure. More advanced techniques are applied to measure intra- and inter-molecular proximities that may be indicative of biologically interesting interactions. Backbone and side chain torsion angles provide local structural constraints. Site-specific relaxation and N-H dipolar coupling measurements are indicative of dynamics in specific residues and domains. T_2 filtering experiments are used to observe site-specific solvent exposure, which sites are proximate to water and which are sequestered away. In a lipid environment, protein-lipid interactions are probed through MAS ssNMR. The orientations and motions of the lipids can be probed through line shapes obtained in static ssNMR. All of these results are combined to give a detailed structural model of the complex biological system at hand. In the following sections and chapters

of this thesis, I apply these concepts to answer complex questions pertaining to structural and dynamic features of polyQ-containing fibrils and Cav1.

1.3 PROBING POLYQ-CONTAINING FIBRIL STRUCTURE AND FORMATION VIA HUNTINGTIN N-TERMINAL FRAGMENT AND POLYQ PEPTIDES

Huntington's disease (HD) is a neurodegenerative disease leading to degradation of cognitive and motor functions and poses a risk for ~200,000 Americans [44]. This disorder is one of at least ten heritable human diseases associated with the expansion of glutamine repeats in the associated protein. HD is caused by an expansion of CAG repeats encoding for a polyglutamine (polyQ) tract in the exon 1 of the N-terminus of the huntingtin (htt) protein. A polyQ tract longer than 35 Gln residues is considered pathological. Although these diseases are associated with the self-aggregation of the polyQ domain to form fibril aggregates, the nature of the most toxic species is unknown. Determining structural features of the mature aggregates is a crucial step toward understanding the misfolding and aggregation pathway.

The htt exon 1 consists of a 17-residue mixed hydrophobic and hydrophilic N-terminus (htt^{NT}), a polyQ tract of variable length, and a proline-rich domain (PRD) that includes two polyproline stretches of 11- and 10-residues. To probe site-specific structural and motional features of the polyQ amyloid core and its flanking domains in htt exon 1, I have used MAS ssNMR on specifically isotopically labeled peptide fibrils, including shortened htt exon 1 peptides, model polyQ peptides, and finally htt exon 1. **Sections 1.3.1 and 1.3.2** contain modified sections from the following peer-reviewed publications:

1) Sivanandam, V. N., Jayaraman, M., **Hoop, C.L.**, Kodali, R., Wetzel, R., and van der Wel, P.C.A. (2011) The aggregation-enhancing huntingtin N-terminus is helical in amyloid fibrils, *J Am Chem Soc 133*, 4558-4566.

2) Mishra, R., **Hoop, C.L.**, Kodali, R., Sahoo, B., van der Wel, P.C.A., and Wetzel, R. (2012) Serine phosphorylation suppresses huntingtin amyloid accumulation by altering protein aggregation properties, *J Mol Biol. 424 (1-2)*, 1-14.
3) Kar, K., **Hoop, C.L.**, Drombosky, K.W., Baker, M.A., Kodali, R., Arduini, I., van der Wel, P.C.A., Horne, W.S., Wetzel, R. (2013) β-hairpin-mediated nucleation of polyglutamine amyloid formation, *J Mol Biol. 425 (7)*, 1183-97.

1.3.1 Using Huntingtin N-Terminal Fragments to Explore Structure and Dynamics of Huntingtin Fibrils

Model peptides were used to study aggregation properties of htt exon 1 fibrils, which incorporated the critical domains of the htt exon 1, referred to as huntingtin N-terminal fragments (HNTF). These peptides contain the htt^{NT} (sequence: MATLEKLMKAFESLKSF), a 30-residue-long polyQ tract, and a ten-residue polyproline domain. Use of synthetic peptides allows for site-specific isotopic labeling for interpretation of NMR data. Strategically placed isotopic labels, allow us to characterize structural and dynamic features of mature fibrils. Fibril preparation has been reproducible from the synthetic peptides and gave sufficient yields. They serve as a valuable model system for htt exon 1, as chemical shifts, and thus structure, of the amyloid core are identical and little change is observed in aggregation behavior and kinetics between htt exon 1 and HNTF peptides and fibrils [45, 46]. Together with Dr. Sivanandam, who did early work on HNTF studies by ssNMR, I have used ssNMR to provide a number of key insights into the structure and stability of the htt^{NT} and polyQ in mature HNTF fibrils.

As the initiation of aggregation is a critical event, the dramatic effects of certain flanking sequences next to the polyQ have attracted much attention, with a particular focus being on the htt^{NT}. The presence of this short segment severely accelerates the aggregation process [47-49].

To gain site-specific structural information on the htt^{NT} and the polyQ domain, three differently labeled peptides were prepared for MAS ssNMR featuring U-^{13}C,^{15}N-labeled residues in selected positions: Ala2, Leu7, Phe17; Ala10, Phe11, Leu14, Gln18; and Leu4, Lys6, Ser16, Gln19. Two-dimensional ^{13}C-^{13}C DARR experiments [35] were used to assign the ^{13}C resonances in the mature fibrils (**Figure 1.8**).

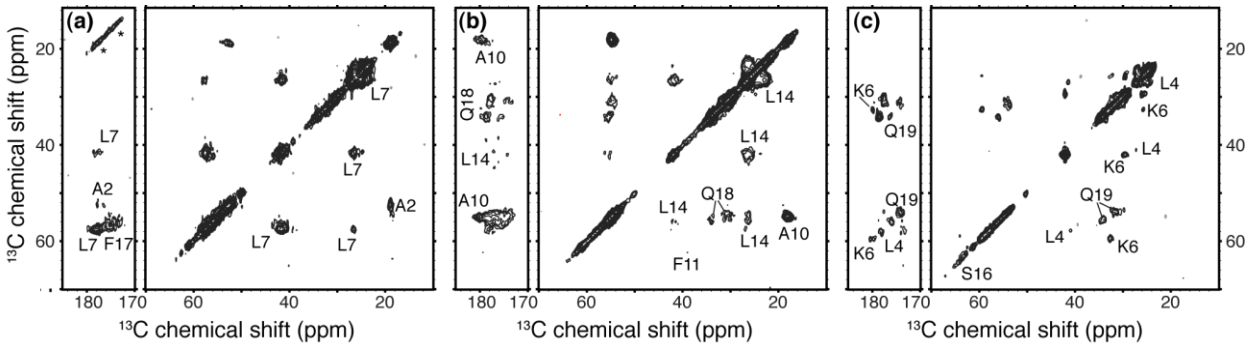

Figure 1.9 Assignment of ^{13}C resonances in isotopically labeled residues.

2D ^{13}C-^{13}C ssNMR spectra using 25 ms DARR mixing, providing 1-2 bond transfers. Measurements on site-specific isotopically labeled HNTF samples are shown in panels (a)-(c), with (a) obtained at 800MHz ^{1}H frequency and 16 kHz MAS, and (b, c) at 600 MHz ^{1}H frequency and 10 and 13 kHz MAS, respectively. Spinning side bands are marked with asterisks. For each spectrum aliphatic-to-carbonyl (left) as well as the intra-aliphatic (right) spectral regions are shown. Reprinted with permission from Sivanandam, V. N., Jayaraman, M., Hoop, C.L., Kodali, R., Wetzel, R., and van der Wel, P.C.A. The aggregation-enhancing huntingtin N-terminus is helical in amyloid fibrils, J Am Chem Soc 133, 4558-4566. Copyright (2011) American Chemical Society.

The assigned chemical shifts are telling of secondary structure as discussed in **Section 1.2.1**. **Figure 1.10** shows a graphical representation of the secondary shifts of the labeled

residues, calculated through subtraction of random coil chemical shifts [50]. The derived secondary structure elements are mapped onto the primary structure in **Figure 1.10b**.

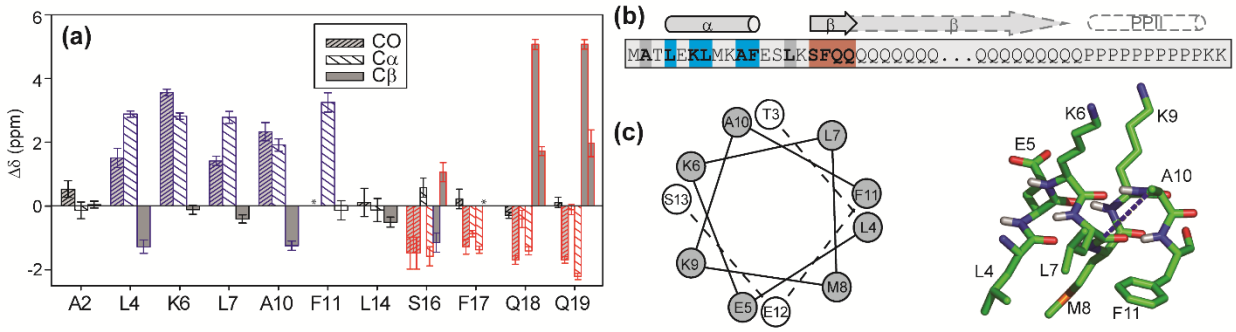

Figure 1.10 Site-specific secondary structure assignment of htt^{NT}Q$_{30}$P$_{10}$K$_2$ fibrils.

(a) Overview of secondary shifts (Δδ) of the labeled C', Cα and Cβ sites, revealing helicity (blue) in residues spanning the residues 4-11, whereas β-sheet structure (red) is seen in residues 16-19. Residues 2 and 14 match neither β-sheet nor α-helical conformations (black). Asterisks mark unobserved labeled sites. Bars are split for sites where two conformations were detected (based on doubled chemical shifts). (b) Schematic representation of observed secondary structure alongside the primary sequence (top), with the labeled residues in bold and color-coded as in (a). Bottom: helical wheel view of the helical segment (observed helicity spans the residues in grey) and a schematic illustration of the amphipathic helix, showing the distribution of hydrophobic residues (front and down) and charged residues (upward) with arbitrary side-chain conformations (prepared using PyMOL, Schrödinger, LLC.). Adapted with permission from Sivanandam, V. N., Jayaraman, M., Hoop, C.L., Kodali, R., Wetzel, R., and van der Wel, P.C.A. The aggregation-enhancing huntingtin N-terminus is helical in amyloid fibrils, J Am Chem Soc 133, 4558-4566. Copyright (2011) American Chemical Society.

Residues 4, 6, 7, 10, and 11 are α-helical in structure. This is suggestive of an amphipathic helix, as illustrated in **Figure 1.10c**. Residues 16-19 are predominantly in a β-sheet conformation, with the exception that one of the two Ser16 forms is neither clearly β-sheet nor α-helical. Residues Ala2 and Leu14 also lack a defined secondary structure, thereby delimiting the length of the helical segment within the htt^{NT} domain. The existence of a contiguous helix was demonstrated

by the observation of an inter-residue (i → i+3) contact between $^{13}C\beta$ of Ala10 and ^{13}CO of Leu7. We also identified the background ^{13}C signals due to the unlabeled prolines in the polyPro segment and found them to be similar to those reported for a PPII helical conformation [51].

From kinetics studies, it is clear that the htt^{NT} plays a critical role in accelerating the aggregation of HNTF fibrils [48, 49, 52]. It is likely that the htt^{NT} α-helices interact with each other to bring about a high local concentration of polyQ, thereby accelerating polyQ self-aggregation, which drives fibril maturation.

In mouse models, that mutation of Ser13 and Ser16 to Asp in the htt^{NT} has been shown to abolish fibril aggregation and neurodegeneration [53]. These sites are at the edges of the α-helical and the β-sheet regions observed by ssNMR in fibrils. To probe for structural changes that may contribute to this reduction in toxicity, we replaced Ser13 and Ser16 with Asp in an HNTF peptide (S13D/S16D mutant HNTF). Aggregation of S13D/S16D HNTF is indeed slowed compared to the WT. The mutant fibrils have a more fractured morphology as observed by TEM. These fibrils were shown to have higher accessibility to trypsin cleavage. All of these measures point to a decrease in stability of the S13D/S16D HNTF fibrils. Any secondary structural changes contributing to the decreased toxicity should be apparent through chemical shift changes in ssNMR. We selectively isotopically labeled residues spanning the htt^{NT} and the first Gln in S13D/S16D HNTF that were previously labeled in the wild-type peptide: Lys6, Ala10, Leu14, and Gln18.

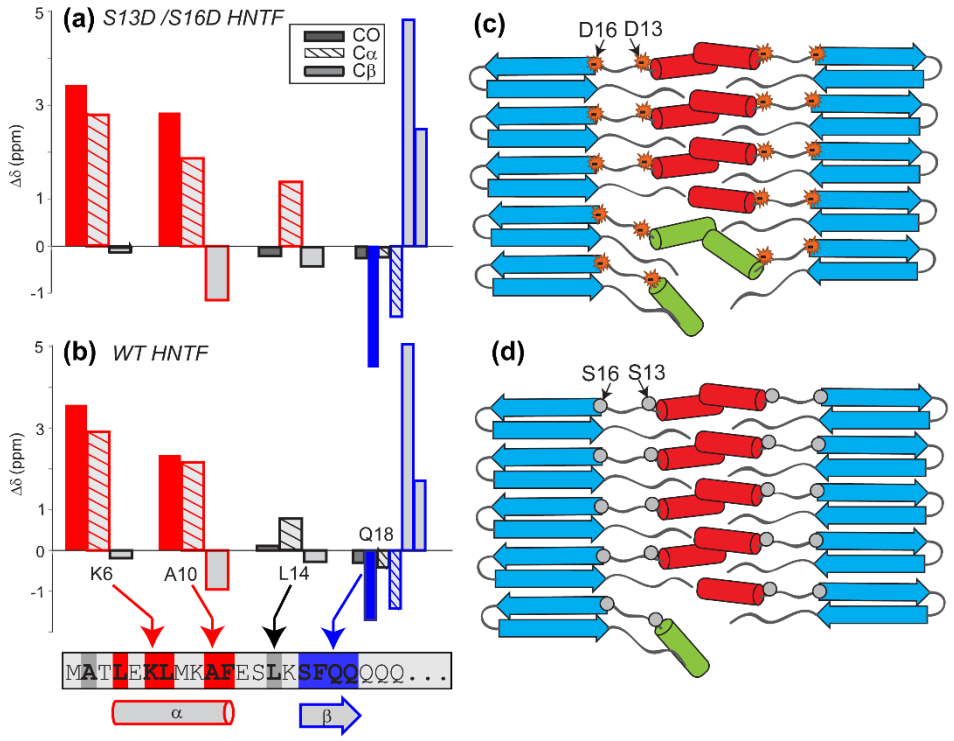

Figure 1.11 Structural consequences of Ser modification in mature htt^{NT}Q$_{30}$P$_{10}$K$_2$ aggregates.

Graphical representation of the ^{13}C secondary chemical shifts ($\Delta\delta$) obtained for indicated isotopically labeled residues based on ssNMR of mature aggregates from (a) S13D/S16D HNTF and (b) WT HNTF peptides. The local secondary structure inferred from those values is shown. Red bars indicate α-helicity, whereas blue bars represent β-sheet structure. Black bars indicate a lack of defined secondary structure. Split bars for Gln18 indicate two sets of NMR signals associated with two conformations. (c,d) A model for the htt exon 1-like fibril structure, integrating observations from ssNMR, EM, and trypsin cleavage. PolyQ β-sheets are in blue, more rigid htt^{NT} α-helices are in red, and more mobile htt^{NT} α-helices are in green. Positions 13 and 16 are in a segment lacking secondary structure [21]. The figure shows how increased mobility (green helices) caused by charge repulsion (c) leads to greater exposure of htt^{NT} sequences to trypsin access compared to more stably packed WT htt^{NT} (d). In both WT and mutated HNTFs, the bulk of the htt^{NT} helices (red) remain stably involved in structure and detectable by MAS ssNMR. Adapted from Mishra, R., Hoop, C.L., Kodali, R., Sahoo, B., van der Wel, P.C.A., and Wetzel, R. Serine phosphorylation suppresses huntingtin amyloid accumulation by altering protein aggregation properties, *J Mol Biol. 424 (1-2), 1-14, Copyright (2012), with permission from Elsevier.*

Surprisingly, no significant chemical shift change was observed in any of the labeled residues (**Figure 1.11a and b**). Despite the lack of structural change, the decrease in stability of the mutant fibrils seems to be correlated with the reduction in toxicity. The negatively charged Asp residues may repel one another, and break up interactions between the htt^{NT} α-helices (**Figure 1.11c and d**). As interactions among the htt^{NT} α-helices accelerates aggregation, diminishing these interactions would decrease the self-aggregation of polyQ and therefore fibril formation. Even though the htt^{NT} α-helices pack less stably in the S13D/S16D HNTF mutant, the signal in MAS ssNMR is dominated by the rigid components, which have the same secondary structure as in the WT HNTF fibrils. Therefore, the Ser-Asp mutations do not change the secondary structure within the mature fibrils but rather decrease the stability of the htt^{NT}-htt^{NT} interactions and are correlated with decreased toxicity.

1.3.2 PolyQ Amyloid Core Has a Unique ssNMR Signature

Most of the htt^{NT} labeled sites (Ala2, Leu4, Lys6, Leu7, Ala10, Phe11, Leu14) in both S13D/S16D and WT fibrils gave single resonances. However, multiple signals were observed in the polyQ domain (Gln18 and Gln19) and residues just prior (Ser16 and Phe17). Intriguingly, these doubled sets of resonances are nearly identical between Gln18 and Gln19 (**Figure 1.12a**). The same sets of chemical shifts are also observed in simple polyQ lacking htt exon 1 flanking sequences. $K_2Q_{30}K_2$ fibrils labeled in only a single Gln display an identical chemical shift pattern. In unlabeled $K_2Q_{31}K_2$ fibrils, the bulk Gln gives the same signals. We refer to the characteristic chemical shifts repeatedly observed in the polyQ core as the "signature motif." The two conformers are most identifiable by the differences in their close Cβ and Cγ chemical shifts,

where ‘form a’ has Cβ and Cγ chemical shifts near 34ppm and those of ‘form b’ are near 31ppm. Doubling of chemical shifts reflects coexistence of multiple conformations, either attributed to differences between polymorphic fibrils (as discussed in **Section 1.1.1**) or structural variations within a single type of fibril. These conformations seem to be characteristic of Glns of the polyQ amyloid core.

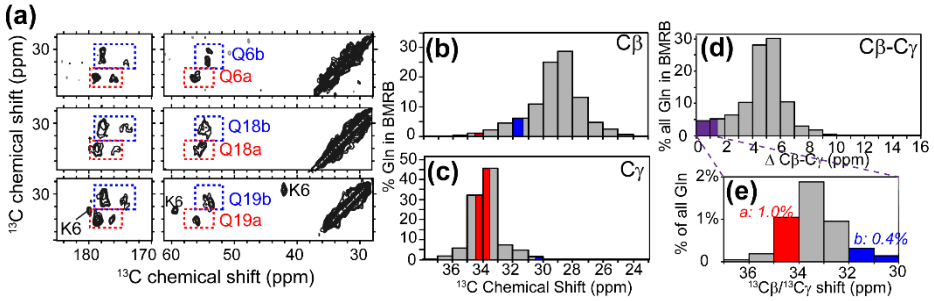

Figure 1.12 Doubling and similarity of glutamine resonances.

(a) 2D ^{13}C-^{13}C DARR spectral regions highlighting the Gln18 (middle panel) and Gln19 resonances (bottom panel) in htt$^{NT}Q_{30}P_{10}K_2$. For each residue we observe two distinct sets of resonances, marked as ‘a’ (red) and ‘b’ (blue). The cross-peak patterns suggest nearly identical chemical shifts for the Cβ and Cγ sites in each case. The top panel shows the virtually identical chemical shifts of a singly labeled glutamine within $K_2Q_{30}K_2$ fibrils (U-^{13}C,^{15}N-Q6 - the fourth residue within the polyQ). (b-e) For the >10,000 Gln in the Biological Magnetic Resonance Bank (BMRB) database (bar graphs), the Gln Cβ (b) and Cγ shifts (c) are generally well separated. In each of the two polyQ conformers, form a (red bars) and form b (blue), there is instead only a small shift difference between Cβ and Cγ. The Cβ shift of form a is highly unusual, and the Cβ and Cγ shifts of form b are both atypical. (d) The difference in Cβ and Cγ chemical shifts plotted as the percentage of all Gln in the BMRB that have that difference. (e) The average shifts of the few BMRB Gln that feature Cβ/Cγ shifts within 1 ppm from each other. The particular shifts of conformer a (red) or b (blue) are seen in 1.0% and 0.4% of all reported Gln, respectively. Panel (a) adapted with permission from Sivanandam, V. N., Jayaraman, M., Hoop, C.L., Kodali, R., Wetzel, R., and van der Wel, P.C.A. The aggregation-enhancing huntingtin N-terminus is helical in amyloid fibrils, J Am Chem Soc 133, 4558-4566. Copyright (2011) American Chemical Society.

To examine the likelihood that these chemical shifts are reproduced by chance, without preservation of structure, we extracted the ^{13}C shifts of >10,000 Gln residues reported in the Biological Magnetic Resonance Data Bank (BMRB; http://www.bmrb.wisc.edu) of proteins studied by NMR. Protein NMR signals for the Cβ and Cγ of Gln are typically well separated in chemical shift (**Figure 1.12b-d**). In contrast, in each of the two observed polyQ Gln conformers, 'form a' (indicated in red) and 'form b' (marked in blue), there is an unusually small shift difference between Cβ and Cγ. **Figure 1.12e** summarizes the (small) subset of Gln in the BMRB that do have a similarly small Cβ/Cγ shift difference, and reveals that only 1.0% and 0.4%, respectively of Gln match either of the observed chemical shift values. The unlikeliness of the observed shifts is also apparent from inspection of the actual shift values in **Figure 1.12b and c**, as the Cβ shift of 'form a' (red) is in itself very unusual, and both the Cβ and Cγ shifts of 'form b' are atypical. Thus, the consistent and reproducible observation of *both* sets of resonances in polyQ amyloids (here and elsewhere [21, 22, 54]) seems exceedingly unlikely to occur by chance. Further studies of this seemingly unique polyQ structure are necessary to fully characterize the nature of the two co-existing forms. Further studies of the two distinct Gln conformers and their incorporation into the complex amyloid-like fibril structure are presented in **Chapter 3.0** .

We have now observed this doubling of amyloid core chemical shifts in fibrils of model peptides in context of HNTFs and polyQ. PolyQ model peptides have been useful in understanding the origin of the two major core Gln species. Mechanistic studies by the Wetzel group probing the role of β-hairpins in aggregation and fibrils relied on PG-based mutations [24, 25]. To encourage β-hairpin formation in polyQ peptides, which was found to lead to enhanced aggregation, a Pro-Gly motif was incorporated in the middle of the Gln stretch ($K_2Q_{11}PGQ_{11}D_2$),

and we conducted MAS ssNMR experiments on these fibrils. As shown in **Figure 1.12**, reports from our labs and others [21, 22] have established that Gln residues within the core of normal polyQ fibrils feature two sets of reproducible NMR signatures. Such ^{13}C chemical shifts are sensitive to the local structure as well as dynamics [55], and are therefore sensitive probes of the structure within the amyloid fibril core. Indeed, amyloid fibrils in different polymorphic forms (as discussed in **Section 1.1.1**) are often identified based on their chemical shifts, which in turn reflect variations in their internal structure [56]. To probe for any changes in the amyloid core structure in mutant fibrils, we introduced a specifically ^{13}C,^{15}N-labeled Gln into the peptide $K_2Q_{11}PGQ_{11}D_2$. The label was placed in the 8th Q in the first Q_{11} segment (i.e., Q10). Again, this single labeled Gln residue yields a doubled set of resonances, indicating two distinct conformations. These are the same doubled resonances, populated in approximately equal amounts, previously found in amyloid fibrils of both simple polyQ and the polyQ portion of HNTFs shown in **Figure 1.13** [21, 22]. For comparison, the spectrum of $K_2Q_{30}K_2$, with the fourth Gln of the polyQ track isotopically labeled, is also shown (**Figure 1.13b**). We concluded that the incorporation of these β-hairpin-promoting mutations does not appreciably modulate these key indicators of the internal amyloid core structure.

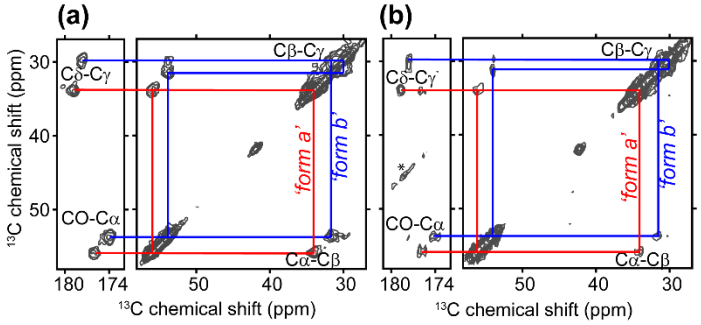

Figure 1.13 Doubled Gln chemical shifts are consistent between simple polyQ and polyQ incorporating a Pro-Gly motif.

(a) Doubled cross-peak pattern of a single labeled Gln in mature fibrils prepared from [U-^{13}C,^{15}N-Q10] $K_2Q_{11}PGQ_{11}D_2$ reproduce the peaks in fibrils from (b) [U-^{13}C,^{15}N-Q6] $K_2Q_{30}K_2$ peptide that lacks β-hairpin encouraging motifs. Red and blue lines mark peaks of the two distinct conformers. Asterisks mark spinning side bands. Spectra were obtained at 600 Mz (^1H), using 10 kHz MAS and 8ms or 25ms DARR mixing, respectively. Adapted from Kar, K., Hoop, C.L., Drombosky, K.W., Baker, M.A., Kodali, R., Arduini, I., van der Wel, P.C.A., Horne, W.S., Wetzel, R. β-hairpin-mediated nucleation of polyglutamine amyloid formation, J Mol Biol. 425 (7), 1183-97. Copyright (2013), with permission from Elsevier.

1.3.3 Summary and Outstanding Questions

Through strategic isotopic labeling of specific residues in the htt^{NT} and polyQ domains of HNTFs, we identified a mobile α-helix in the htt^{NT}. This α-helix is maintained in S13D/S16D mutant HNTFs, models of htt exon 1 mutants that were shown to abolish fibril formation in mice models. The htt^{NT} appears to initiate and stabilize fibril formation through intermolecular interactions between htt^{NT} α-helices. An EPR study of htt exon 1 supports tertiary and quaternary contacts between htt^{NT} α-helices in mature fibrils [23]. Destabilizing the interactions between α-helices greatly reduces fibril formation. Outstanding questions remain on the impact of the flanking sequences, both the N-terminal htt^{NT} and C-terminal proline-rich domain (PRD), on the polyQ core. How is the localization of the β-structured amyloid core affected? Exactly how mobile is the htt^{NT} α-helix in the fibrils, given its stabilizing role in the fibrils? Where is the transition from β-strand to PPII structure in the fibrils? These questions are addressed in **Chapter 2.0** .

In studies of the polyQ core in fibrils of both HNTFs and polyQ peptides, we have observed a doubling of chemical shifts that are very unique and seemingly characteristic of Gln residues in the amyloid core. As chemical shifts are indicative of structure, this brings to question, what are the characteristic structures and dynamic features of the two Gln conformers? What is the origin of the "signature motif"? How are the two conformers incorporated into the amyloid fibril assembly? These questions are addressed in **Chapter 3.0** .

1.4 SUMMARY OF PROJECTS

The flanking segments of the polyQ domain in htt have different effects on the rate of aggregation. The C-terminal proline-rich domain inhibits aggregation by modulating the structural ensemble of the monomeric state, through induction of PPII structure into a portion of the polyQ domain. The α-helical N-terminal htt^{NT} accelerates aggregation, stabilizing oligomers and fibrils. In **Chapter 2.0** , I show that HNTF fibrils are structurally similar to htt exon 1 fibrils as observed by NMR, making them useful models of htt exon 1. I discuss how we used ssNMR of the mature HNTF fibrils to probe the influence of the flanking domains on the polyQ tract, in particular the boundaries of the amyloid core. We found that the β-structured amyloid core extends to the penultimate Gln. We also showed that the htt^{NT} α-helix experiences solvent-coupled dynamics, consistent with a molten-globule-like structure.

In various polyQ-containing fibrils, the Gln residues of the polyQ amyloid core have consistently shown a characteristic doubling of NMR chemical shifts, telling of an equal distribution of two distinct Gln conformers (as shown in **Figure 1.12** in **Section 1.3.2**). The chemical shift patterns observed seem unique to polyQ. In **Chapter 3.0** , I discuss how we used MAS ssNMR to distinguish the two conformers in structure and dynamics, in part via advanced ssNMR methods that constrain backbone and side chain torsion angles. Additionally, we used ssNMR to probe the impact of incorporation of Pro-Gly motifs into polyQ on the fibril structure. These studies complement mechanistic studies of polyQ aggregation and shed light on the origins of the doubled signature of the polyQ amyloid core. Combined, the results allow us to

present a preliminary new model of the polyQ amyloid core and reveal key features of its assembly.

The CSD of Cav1 has several critical roles, including protein recognition, oligomerization, and cholesterol binding. In **Chapter 4.0** , I discuss how we have employed several spectroscopic methods, including FTIR, CD, and ssNMR to probe the structure of the CSD in Cav1 fragments bound to cholesterol-rich membranes. We probed the structure of residues within the putative CRAC motif and residues involved in protein binding within the CSD. We have observed β-structure in CSD, contrasting with earlier reports, yet consistent with secondary structure predictions based on preference by primary sequence. This β-structure seems consistent with general features of cholesterol-binding domains and their function in protein recognition and oligomerization.

2.0 POLYGLUTAMINE AMYLOID CORE BOUNDARIES AND FLANKING DOMAIN DYNAMICS IN HUNTINGTIN FRAGMENT FIBRILS DETERMINED BY SOLID-STATE NMR

Adapted with permission from**: Hoop, C.L.**, Lin, H.K., Karunakar, K., Zhipeng, H., Poirier, M.A., Wetzel, R., and van der Wel, P.C.A. Polyglutamine amyloid core boundaries and flanking domain dynamics in huntingtin fragment fibrils determined by solid-state NMR, *Biochemistry 53 (42)*, 6653-66. Copyright (2014) American Chemical Society.

Author contributions: C.L.H. acquired and analyzed the MAS ssNMR data. H.K.L. expressed, purified, and fibrillized [U-^{13}C,^{15}N]-htt exon 1 and acquired its MAS ssNMR spectra. K.K. purified and fibrillized all synthetic huntingtin N-terminal fragment and polyQ peptides. H.Z. and M.A.P. provided the htt exon 1 construct. C.L.H., H.K.L., K.K., R.W., and P.v.d.W. wrote the publication.

In HD, expansion of a polyQ domain in the htt protein leads to misfolding and aggregation. There is much interest in the molecular features that distinguish monomeric, oligomeric, and fibrillar species that populate the aggregation pathway and likely differ in cytotoxicity. The mechanism and rate of aggregation are greatly affected by the domains flanking the polyQ segment of htt. A “protective” proline-rich flanking domain inhibits

aggregation by inducing polyproline II structure (PPII) within an extended portion of the polyQ. The N-terminal flanking segment (htt^{NT}) adopts an α-helical structure as it drives aggregation, helps stabilize oligomers and fibrils, and is seemingly integral to their supramolecular assembly. Here, we use ssNMR to probe how, in the mature fibrils, the htt flanking domains impact the polyQ domain, and in particular the localization of the β-structured amyloid core. Using residue-specific labeling, we find that the amyloid core ends just prior to the prolines. We probe the structural and dynamical features of the remarkably compact β-sheet to PPII transition, and discuss the potential connections to certain htt-binding proteins. We also examine the N-terminal flanking domain outside the polyQ amyloid core. Despite presumed structural and demonstrated stabilizing roles of the htt^{NT} α-helix in the fibrils, quantitative ssNMR measurements of residue-specific dynamics show that it experiences distinct solvent-coupled dynamics. This dynamical feature seems reminiscent of molten-globule-like α-helix-rich features attributed to the non-fibrillar oligomeric species of various amyloidogenic proteins.

2.1 INTRODUCTION

HD is one of several human disorders caused by a genetic mutation that expands CAG repeats within genes encoding for several different proteins [57]. Expansion of the polyglutamine tracts in these proteins beyond a disease-specific threshold correlates to a decrease in the age of pathological onset and an increase in toxicity [58]. In HD and other polyQ expansion disorders, disease pathology is associated with protein misfolding and aggregation, including the formation of amyloid-like fibrils. Some experimental evidence supports a toxic role of the polyQ amyloid-like aggregates [59-62], with toxicity dependent on the nature of the

aggregates [62], but others argue for a protective, non-toxic role for the mature aggregates [63]. Given the potential impact on disease toxicity and the fact that both polyQ expansion and some level of misfolding, aggregation, or oligomerization are common features among the CAG repeat diseases, much work has been done to characterize the aggregation process. The aggregation propensity of polyQ model peptides increases with Gln repeat length [64]. On top of this, variations in the flanking sequences can have a dramatic impact on the conformation and aggregation propensity of the polyQ tract and thus may in part explain threshold differences between the CAG expansions disorders [47, 49 , 65-67].

In HD, Gln expansion occurs in the htt protein leading to protein misfolding and aggregation. Observed aggregates contain N-terminal fragments of htt [68] generated by protease activity [69] or aberrant RNA splicing [70]. In animal studies, overexpression of the fragment corresponding to mutant htt's exon 1 leads to neurodegeneration that mimics key features of HD [71]. On either side of the polyQ domain, htt exon 1 features flanking segments that act as binding sites for htt-binding proteins and modulate the binding to membranes and subcellular targeting [47, 72, 73]. In cellular studies of mutant htt exon 1, the flanking domains modulate not only the aggregation propensity, but also aggregate morphology, recruitment of other proteins, and toxicity [72, 74-76]. A Pro-rich domain (PRD) that follows the polyQ domain greatly reduces aggregation propensity [47, 65-67, 77], leading to suggestions that it may in part function to reduce the risk of misfolding and disease [75, 78]. Preceding the polyQ domain, htt features a 17-residue sequence of mixed hydrophobic and hydrophilic amino acids, referred to as NT17 or htt^{NT}, which in contrast to the PRD, initiates and accelerates aggregation [49]. In doing so, htt^{NT} enables a distinct multi-stage aggregation pathway that features α-helix-rich oligomeric species (in contrast to polyQ model peptide aggregation) [21, 54, 79].

Disease toxicity in HD is thought to reflect a gain of toxic function that accompanies conformational changes in the mutant protein. As in other protein misfolding diseases, different conformational states populate the misfolding and aggregation pathway. These different species could all contribute to disease toxicity in their own way, whether by being toxic in their own right, causing toxicity by sequestering vital polyQ-containing proteins, or reducing toxicity by sequestering toxic oligomers or monomers. Therefore, there has been much interest in understanding the conformational changes that occur during aggregation in HD and other deposition-related disorders.

As for other amyloidogenic proteins, nucleation of a β-sheet-rich amyloid-like conformation is a pivotal event in the aggregation of mutant htt N-terminal fragments. This nucleation event can take place in the monomeric or oligomeric state and precedes and enables the formation of anti-parallel β-sheet-rich fibrillar aggregates that have all of the features of amyloid-like fibrils. Earlier studies have shown that the C-terminal PRD counteracts the nucleation of β-sheet structure, possibly by inducing polyproline II (PPII) structure in an extended stretch of the preceding polyQ segment [65-67, 80, 81]. These observations are primarily based on solution studies of various monomeric species, and reflect a backbone-mediated effect of the PPII-structured PRD on the conformational ensemble of the non-aggregated polyQ domain. It has remained unclear to what degree this inferred induction of PPII in the glutamines persists in the oligomeric or fibrillar species. One might expect that also in the fibrils one or more Gln at the C-terminal end of the polyQ domain remain outside the amyloid core, perhaps in a PPII conformation.

The N-terminal flanking domain, htt^{NT}, is responsible for the formation of oligomeric assemblies [48, 49, 79, 82]. The oligomerization precedes the nucleation of β-sheet structure and

is instead characterized by extensive α-helical structure [21, 54, 79]. This α-helicity is largely attributed to the htt^{NT} segment although it may also extend into the polyQ domain [83, 84], since the latter has a length-dependent propensity for α-helical structure that is modulated by flanking domains [2, 65, 83-89]. The oligomer ensembles seem to have an enhanced propensity (compared with the monomers) for β-sheet nucleation, translating into greatly accelerated formation of the β-sheet-rich fibrils [48, 49]. Thus, the htt^{NT} segment effectively counteracts the effects of the PRD, and may allow aggregation under conditions where the PRD alone would effectively prevent fibril formation [49]. Also in the fibrils, intermolecular interactions among the htt^{NT} amphipathic α-helices appear to contribute to the stability and supramolecular structure [23, 54]. Weakening of such interactions, e.g by htt^{NT} phosphorylation, results in reductions in aggregation propensity, aggregate stability, and disease toxicity [53, 54].

We previously showed by MAS ssNMR and transmission electron microscopy (TEM) that htt peptide fragments aggregate into fibrils with a β-sheet core, but with different secondary structures in the flanking domains (see Section 1.3.1) [21]. Subsequently, EPR studies of htt exon 1 outfitted with spin labels suggested a similar domain assembly of mutant exon 1 fibrils, while also yielding evidence of fibrillar htt^{NT}-htt^{NT} interactions [23]. Our study of fibril formation in presence of phosphomimetic htt^{NT} mutations also indicated a role for intermolecular htt^{NT} interactions in the stabilization of fibrillar and oligomeric assemblies (See **Section 1.3.1**) [54], with apparent consequences for cellular toxicity [53, 54]. The connection between aggregate structure, stability and toxicity [62] may in part reflect changes in the access to proteases and chaperones as they interact with oligomeric and aggregated states of htt [48, 54].

Due to the limited structural data on the fibrils and in particular the oligomeric species, we lack an in-depth understanding of the structural transformations within mutant htt during

misfolding and self-assembly. For instance, it is uncertain whether or how the PRD affects the secondary structure of the C-terminal tail end of the polyQ domain in the fibrils. Are the preceding Gln in a PPII-like structure, as they seem to be in the monomers, or is this "protective" effect overcome once nucleation and propagation of β-structure has taken place? How accessible and dynamic are the flanking domains in mature fibrils, given their apparent role in stabilizing the various aggregated species? Here, we address these questions through the use of MAS ssNMR on fibrils obtained from a peptide construct ($htt^{NT}Q_{30}P_{10}K_2$) that aggregates into fibrils [21]. These constructs recapitulate the aggregation behavior of htt exon 1 [46] and have been used in biophysical, mechanistic and structural studies by us as well as others [21, 49, 53, 54, 90-94]. This includes several detailed studies of the individual and combined impacts of the flanking domains on aggregation kinetics and aggregate morphology [21, 49, 65, 91]. These synthetic constructs facilitate the residue-specific labeling that permits unambiguous identification of specific residues within the various domains. This in turn, allows for a careful analysis of these residues' structural and dynamic features. We complement these residue-specific insights with the first ssNMR studies of fibrils prepared from uniformly labeled, full-length mutant htt exon 1. We observe by MAS ssNMR that, in contrast to its reported effect on the monomeric species, the proximity of the oligoPro flanking domain does not prevent β-sheet structure near the C-terminal end of the polyQ domain. Fibril maturation leads to formation of a relatively well-ordered amyloid core that consists of β-sheet structure and contains most of the polyQ domain. Quantitative site-specific dynamics measurements reveal that, despite their stabilizing role and supramolecular interactions, the htt^{NT} α-helices outside this amyloid core undergo molten-globule-like motions.

2.2 EXPERIMENTAL PROCEDURES

2.2.1 Preparation of Residue-Specifically Labeled Fibrils

Residue-specifically labeled peptides were prepared by solid-phase synthesis at the Small Scale Synthesis facility of the Keck Biotechnology Resource Laboratory of Yale University. Site-specific labeling with U-^{13}C,^{15}N-labeled residues was performed at various positions as indicated in **Table 2.1**. Fmoc- and side-chain-protected ^{13}C,^{15}N-labeled amino acids were purchased from Cambridge Isotope Laboratories (Andover, MA) and Isotec (Sigma-Aldrich, St. Louis, MO). The peptides were obtained crude for in-house purification and disaggregation, as previously described [25, 95]. Aggregation was initiated starting with disaggregated peptide and performed in PBS buffer (pH 7.4) at 37°C, monitored by an HPLC-based sedimentation assay, and allowed to proceed to completion [25, 95]. After being washed with deionized water, mature fibrils were pelleted into 3.2 mm zirconia MAS rotors (Bruker Biospin, Billerica, MA and CortecNet, Voisins-le-Bretonneux, France) by centrifugation using a home-built ultracentrifugal sample packing tool. Samples were kept hydrated and unfrozen unless otherwise stated.

Table 2.1. Nomenclature and site-specific labeling of the $htt^{NT}Q_{30}P_{10}K_2$ peptides.

MAS ssNMR samples contained approximately 8mg of peptide (with the exception of 1.8mg in the LKSQ-labeled peptide sample [21]).

ID	Labeling	Sequence [a]
LQP-labeled	U-^{13}C,^{15}N-[L7, Q47, P48]	MATLEKLMKAFESLKSF-QQQ$_{26}$QQPP$_9$K$_2$
LAQ-labeled	U-^{13}C,^{15}N-[L4, A10, Q46]	MATLEKLMKAFESLKSF-QQQ$_{26}$QQPP$_9$K$_2$
MA-labeled	U-^{13}C,^{15}N-[M8, A10]	MATLEKLMKAFESLKSF-QQQ$_{26}$QQPP$_9$K$_2$
LKSQ-labeled	U-^{13}C,^{15}N-[L4, K6, S16, Q19]	MATLEKLMKAFESLKSF-QQQ$_{26}$QQPP$_9$K$_2$
MF-labeled	U-^{13}C,^{15}N-[M8, F11]	MATLEKLMKAFESLKSF-QQQ$_{26}$QQPP$_9$K$_2$

a) Isotopically labeled residues are underlined.

2.2.2 Preparation of Uniformly Labeled htt Exon 1 Fibrils

I acknowledge Dr. Michelle Poirier for making and providing a maltose binding protein (MBP) construct for full-length htt exon 1. I thank Dr. Hsiang-Kai Lin for expressing the fusion protein in E. coli and preparing fibrils from uniformly ^{13}C,^{15}N-labeled htt exon 1 to complement the residue-specifically labeled fibrils. The employed construct was based on a previously reported construct [96], but features a new linker design that avoids the inclusion of extra residues prior to the htt^{NT} upon cleavage with factor Xa. Overexpression was induced by adding 0.5mM isopropyl β-D thiogalactopyranoside (RPI corp., Mt Prospect, IL), supplementary 0.02%(w/v) ^{13}C-D-Glucose and 0.01%(w/v) ^{15}N-ammonium chloride, and 100 μM $ZnSO_4$. The protein was expressed at 18°C for 16 hrs, after which the cells were harvested by centrifugation and the cell pellets resuspended in PBS (pH7.4) with 1mM phenylmethanesulfonyl fluoride (ACROS, Fair Lawn, NJ). The fusion protein was purified by Histrap FF column (GE

Healthcare) and assessed for purity by SDS-PAGE (12%). Protein identity and isotopic labeling were verified with matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI) time-of-flight (TOF) mass spectrometry at the University of Pittsburgh's Genomics and Proteomics Core Laboratories. Purified fusion protein was buffer-exchanged to PBS in centrifugal filter units (Millipore, Billerica, MA). To remove the MBP fusion domain, release htt exon 1, and induce fibrillation [96], the fusion protein (at 5.8 mg/ml) was treated with factor Xa protease (0.0028 mg/ml) (Promega, Madison, WI) at 37 °C. Generation and identity of the released htt exon 1 were tested by SDS-PAGE (12%) and ESI-TOF mass spectrometry. U-^{13}C,^{15}N-labeled htt exon 1 aggregates were pelleted down, resuspended into 1 mL PBS, and then pelleted into a 3.2 mm MAS rotor using a centrifugal packing tool, as described above.

2.2.3 MAS Solid-State NMR

All experiments were performed with a wide bore Bruker Avance I NMR spectrometer operating at 600 MHz ^{1}H Larmor frequency (14.1 T), using either a standard-bore 3.2 mm HFCN MAS probe or a 3.2 mm EFree HCN MAS probe from Bruker Biospin (Billerica, MA, USA). The sample temperature was controlled using a constant flow of cooled gas. Spectra were acquired with the Bruker Topspin software, but processed and analyzed with NMRPipe, Sparky, and CCPNMR/Analysis software. External referencing to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) was done indirectly via the ^{13}C signals of adamantane [97]. For all spectra shown, additional experimental details can be found in **Appendix A** (**Table A.1**).

2.2.3.1 ^{13}C and ^{15}N Chemical Shift Assignments by MAS ssNMR

Residue-specific ^{13}C and ^{15}N assignments were obtained via 2D ^{13}C-^{13}C and ^{15}N-^{13}C experiments. These experiments involved ^{1}H-^{13}C/^{15}N and ^{15}N-^{13}C cross polarization (CP) steps combined with ^{13}C-^{13}C-transfers via dipolar assisted rotational resonance (DARR) [35] mixing (8-15 ms). Typically 83 kHz two pulse phase modulation (TPPM [98]) ^{1}H decoupling was applied during acquisition and evolution. MAS spinning rates (9-10 kHz) and other details can be found in **Table A.2** in **Appendix A**.

2.2.3.2 Site-Specific Dynamics Measurements by MAS ssNMR

To probe for site-specific dynamics in the htt^{NT} and polyQ domains, ^{15}N longitudinal relaxation was measured for backbone ^{15}N sites using a series of ^{1}H-^{15}N CP experiments incorporating a ^{15}N T_1 relaxation period, analogous to earlier work [99, 100]. These measurements were performed at 19-22 kHz MAS, where systematic MAS-dependent measurements had shown ^{15}N-^{15}N spin diffusion to be effectively suppressed. During acquisition, 83 kHz TPPM ^{1}H decoupling was applied. As a complementary indicator of dynamics, we measured the one-bond N-H dipolar couplings of selected labeled sites. This was done at 10 kHz MAS through use of constant-time DIPSHIFT [101] experiments involving rotor-synchronous incrementation of an $R18_1^7$ symmetry sequence (89 kHz RF power) [102, 103] at the expense of 100 kHz continuous wave ^{1}H decoupling. The experimental data curves were analyzed using numerical simulations with the SPINEVOLUTION program [104], optimizing both the (apparent) N-H distance and exponential damping during the R-sequence mixing time. The apparent N-H distances were converted into dipolar coupling constants (in kHz). As another qualitative measure of dynamics, we performed comparative 1D ^{13}C experiments comparing ^{13}C direct polarization (DP) to signals obtained via dipolar-based ^{1}H-^{13}C

CP, under otherwise identical conditions. These two polarization techniques have differential sensitivity to dynamics, and their comparison has been used to gain insight into local dynamics in various biological systems.

2.2.3.3 Site-Specific Solvent Exposure Measurements by MAS ssNMR

To probe site-specific proximity to the fibril surface and exposure to solvent, we used MAS ssNMR measurements in which protons from highly mobile water molecules are selectively used as a source of polarization. These experiments featured a ^{1}H-^{13}C CP step preceded by a 3 ms T_2 relaxation filter and a variable-time ^{1}H-^{1}H spin diffusion period, using a pulse sequence as previously described [105-107]. These experiments were performed using a range of spin diffusion periods from 0 to 7 ms, after which the re-polarization of the peptide signals was monitored by comparing peak intensities to the un-filtered ^{1}H-^{13}C CP signal. These spectra were obtained at 10 kHz MAS, with 83 kHz TPPM ^{1}H decoupling during acquisition.

2.3 RESULTS

2.3.1 Structural Analysis of the Q/P Junction

Previous studies have shown that within monomeric peptides the C-terminal oligoPro domain disrupts the conformational ensemble of the Pro-proximal Gln residues by inducing PPII structure within a stretch of multiple Gln residues [65, 108-111]. To probe whether or how this disruption of the polyQ domain is retained in fibrils, we performed MAS ssNMR on aggregates prepared from $htt^{NT}Q_{30}P_{10}K_2$ peptides featuring site-specific labeling of residues within and beyond the C-terminal end of the polyQ segment. We previously submitted such aggregates (with differing labeling schemes) to TEM to show their fibrillar morphology [21]. Here we examine fibrils prepared from the LQP-labeled peptide (**Table 2.1**) that is ^{13}C,^{15}N-labeled in Q47 and P48, which constitute the Q/P junction. Using standard 2D ssNMR spectra (e.g. **Figure 2.1a and b**), we obtained assignments for the labeled residues. P48, which is the first Pro of the oligoPro domain, is found to have ^{13}C chemical shifts that match typical chemical shifts of Pro residues in a PPII 3_1 helix [112, 113]. **Figure 2.2** illustrates the nearly perfect correspondence between the P48 peaks (**Figure 2.2a**; black arrows) and the solution NMR chemical shifts (**Figure 2.2c**) of a PPII oligoproline segment in the decapeptide, APSYSPPPPP, bound by α-spectrin SH3 (**Figure 2.2d**; BMRB entry 15013 [112]). We previously [21] found that unlabeled oligoPro in $htt^{NT}Q_{30}P_{10}K_2$ aggregates showed the same shifts (**Figure 2.2b**), indicating PPII helix structure throughout the oligoPro domain. These ssNMR data on fibrils are in agreement with

solution NMR studies of oligoPro in much shorter polyQ-containing peptides in the monomeric state, although no ^{13}C chemical shifts are available for direct comparison [110].

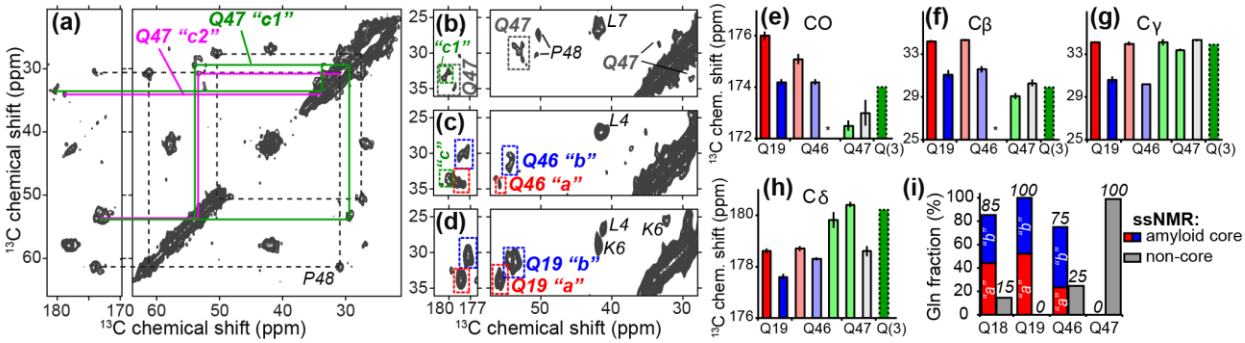

Figure 2.1 ^{13}C chemical shift assignments via 2D MAS ssNMR spectra.

(a) Aliphatic-carbonyl (left) and intra-aliphatic (right) spectral regions from a 8ms DARR ^{13}C-^{13}C 2D spectrum on LQP-labeled htt^{NT}Q$_{30}$P$_{10}$K$_2$ fibrils at 275 K. (b-d) Comparison of individual Gln in different parts of the polyQ domain: (b) Q47 (15 ms DARR, 9.8 kHz MAS), (c) Q46 (8 ms DARR, 10 kHz MAS), and (d) Q19 (25 ms DARR, 13 kHz MAS, data from ref. [21]). In these spectra color-coded lines, boxes and quoted letters ("a", "b", etc) mark the Gln conformers discussed in the text. In (b-d), peaks from labeled residues L4, K6, L7, and P48 are indicated. (e) Backbone and (f-h) side-chain ^{13}C chemical shifts for Gln in htt^{NT}Q$_{30}$P$_{10}$K$_2$ fibrils. Red/blue color-coding indicates conformers "a" and "b" (see refs. [21, 114]). Newly observed resonances for Q46 and Q47 are shown with paler color to indicate the conformers, where applicable. Green conformers for Q46 and Q47 resemble a "Q(3)" conformer reported in polyQ fibrils [115], with a Cδ shift near 180ppm. Missing resonances are indicated with an asterisk. (i) Fractions of residues Q18, Q19, Q46, and Q47 that feature the amyloid core signature, estimated from ssNMR 2D peak volumes.

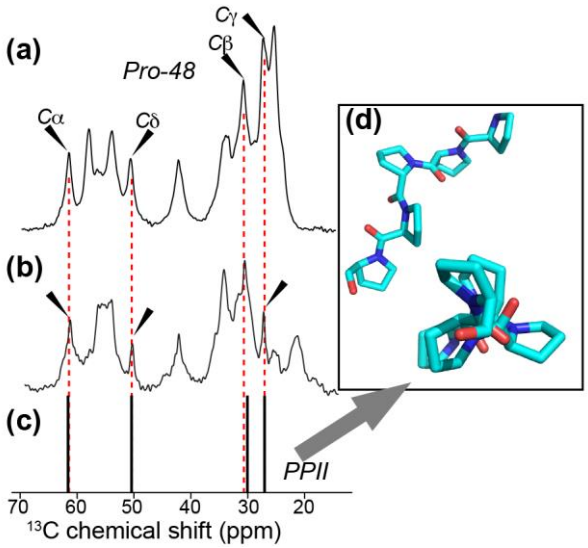

Figure 2.2 ^{13}C chemical shifts for Pro in htt^{NT}Q$_{30}$P$_{10}$K$_2$ fibrils.

(a) ^{1}H-^{13}C 1D CP MAS spectrum of LQP-labeled fibrils at 9.8 kHz MAS, showing the peaks due to the first Pro (P48; black arrows). (b) Natural abundance Pro signals (black arrows) in unlabeled htt^{NT}Q$_{30}$P$_{10}$K$_2$ fibrils at 10 kHz MAS [21]. Unmarked peaks are due to other residues, either isotopically enriched (in (a)) or at natural abundance. Adapted with permission from Sivanandam et al. (2011) J. Am. Chem. Soc. 133 (12), 4558–4566. Copyright 2011 American Chemical Society. (c) Chemical shifts of Pro in an example of PPII-structured oligoPro (shown in d), from previously published solution NMR on a chimera of the peptide APSYSPPPPP and α-spectrin SH3 [112].

Next, we examined the extent to which Gln residues preceding oligoPro are affected by the proximity to the oligoPro segment and the end of the polyQ domain. Previously, we have identified a signature chemical shift pattern that characterizes polyQ amyloid core residues [114]. This pattern features two co-existing equal populations of Gln that have different NMR signals. As such chemical shift differences imply distinct conformations, we have referred to these Gln as conformers "a" and "b". They feature unique Cβ and Cγ chemical shifts discussed in **Section 1.3.2** that are highly uncommon in other protein structures [21, 114], but are thus far seen in all ssNMR studies on polyQ amyloid [21, 54, 114, 115]. **Figure 2.1d** shows a representative 2D MAS ssNMR spectrum of htt^{NT}Q$_{30}$P$_{10}$K$_2$ aggregates[21] in which Q19 (the

second Gln in the polyQ segment) is uniformly isotopically labeled. The sets of peaks corresponding to the two amyloid core conformers "a" and "b" are highlighted with red and blue boxes, as well as matching colored labels "Q19a" and "Q19b". We adopt the same conformer-specific color-coding in subsequent figures. The corresponding sets of ^{13}C chemical shifts are also shown as red and blue bar graphs in **Figure 2.1e**. The peaks for the ^{13}C,^{15}N-labeled Q47 (green and magenta lines in **Figure 2.1a**; dashed boxes in **Figure 2.1b**), directly preceding P48, clearly lack these chemical shifts that identify the β-sheet polyQ core. Q47 does display some structural heterogeneity, with two major conformers that are marked as "c1" and "c2" in **Figure 2.1**. The Q47 Cβ and Cγ chemical shifts near ~30 and ~34 ppm, respectively, are very different from the polyQ core signature and more similar to typical Gln found in globular proteins. Thus, Q47 is structurally heterogeneous and not participating in the polyQ amyloid core. Given the multi-residue effect of the PPII-structured oligoPro flanking sequence in the monomeric state, one may expect a similar conformation for residues further into the polyQ segment. To investigate this, we prepared a fibril sample with isotopic labels in Q46 (alongside labeling of L4 and A10 in the htt^{NT} segment; **Table 2.1**). **Figure 2.1c** shows results from a ^{13}C-^{13}C 2D experiment on this sample, again highlighting in color-coded, dashed boxes the peaks assigned to Q46. Unlike Q47, the majority population of this residue does feature the ^{13}C amyloid core shifts (compare to Q19 in **Figure 2.1d**) and is thus found in the amyloid core [21, 54, 114, 115]. A small fraction (~25%) of Q46 adopts chemical shifts that resemble those of Q47, indicating that in some of the peptide chains this residue is outside the core. This contrasts with Q19, which lacks detectable signals with non-core shift values, but resembles Q18, the first residue of the polyQ domain, which also exhibits a small fraction of Gln outside the amyloid core (**Figure 2.1i**). Thus, in these peptide fibrils we see a highly localized conformational disruption of the

amyloid core at both ends of the polyQ domain, which is most pronounced in the single residue just prior to the oligoPro flanking segment.

2.3.2 Solvent Accessibility and Mobility in the Q/P Junction

Having established that the Q/P junction constitutes a remarkably tight transition between two secondary structure elements, we performed MAS ssNMR experiments to assess the solvent accessibility of this structural motif that acts as a recognition site for certain htt-binding proteins [116-119]. First, we directly probe water accessibility via T_2-filtered ^{13}C MAS ssNMR [105], previously used on other amyloid fibrils and membrane proteins [106, 107, 120]. After eliminating signals from (rigid) protein protons (**Figure 2.3b**, red), time-dependent ^{1}H polarization transfer from mobile water reflects the proximity of specific residues to the fibril surface [120]. We observe a differential repolarization of the residues in the different peptide domains (**Figure 2.3c**). The fastest transfer occurs to L7, which is outside the amyloid core and in the N-terminal α-helix (more on this below). At the Q/P junction, Q47 and P48 both show increased water accessibility compared to the dehydrated amyloid core [21].

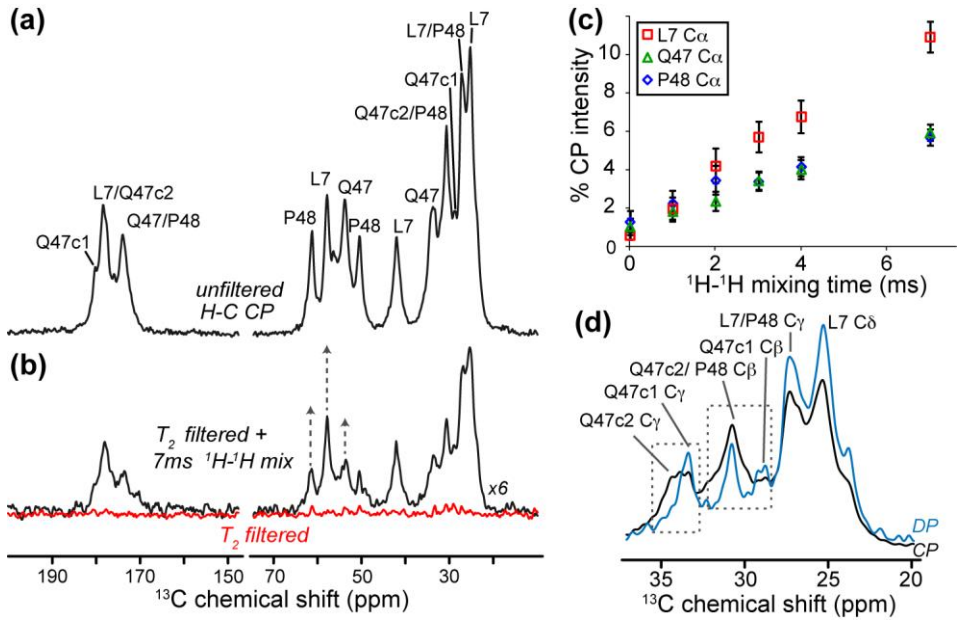

Figure 2.3 Solvent exposure and dynamics at the Q/P junction of htt^{NT}Q$_{30}$P$_{10}$K$_2$ fibrils.

(a) ^{1}H-^{13}C CP spectrum of the LQP-labeled sample. (b) A 3 ms T_2 filter eliminates virtually all signals (red). 7 ms ^{1}H-^{1}H mixing (black) allows ^{1}H polarization transfer from mobile waters into the fibrils. The most prominent peaks are due to L7 in the htt^{NT} helix, followed by P48 and Q47 (each Cα marked with arrow). (c) Polarization build-up curves of Cα sites (marked in (b)) as a function of ^{1}H-^{1}H mixing, normalized relative to the unfiltered ^{1}H-^{13}C CP signal in (a). (d) Comparison of ^{1}H-^{13}C CP (black) and ^{13}C DP (cyan) spectra on LQP-labeled fibrils. Unlike L7, various Q47 peaks are notably narrower in the DP data, reflecting locally increased dynamics. The DP spectrum is rescaled up to enable better comparison of the line widths.

To evaluate local dynamics, we compared the peak intensities of CP-based and DP ^{13}C ssNMR spectra. The dipolar mediated ^{1}H-^{13}C transfer in CP experiments is less effective for mobile sites, whilst the slow relaxation recovery of rigid ^{13}C signals suppresses the most rigid sites in the DP spectra. Thus, mobile and rigid residues can be distinguished, and even spectroscopically "filtered out". At the htt Q/P junction, these experiments reveal populations of Q47 with different levels of dynamics. Looking at the side-chain Cγ (**Figure 2.3d**), the c2 conformer of Q47 is suppressed in the ^{13}C DP experiment (i.e., is rigid), whilst conformer c1 is

present and even reveals a quite narrow linewidth (i.e., slower T_2 relaxation due to increased flexibility). The absolute peak intensities (and their differences between the CP and DP experiments) for both these Q47 conformers are indicative of significantly greater dynamics than the amyloid core (**Figure 2.4**). Thus, Q47 features increased mobility and solvent exposure, consistent with its location outside, or at the surface of, the polyQ amyloid core.

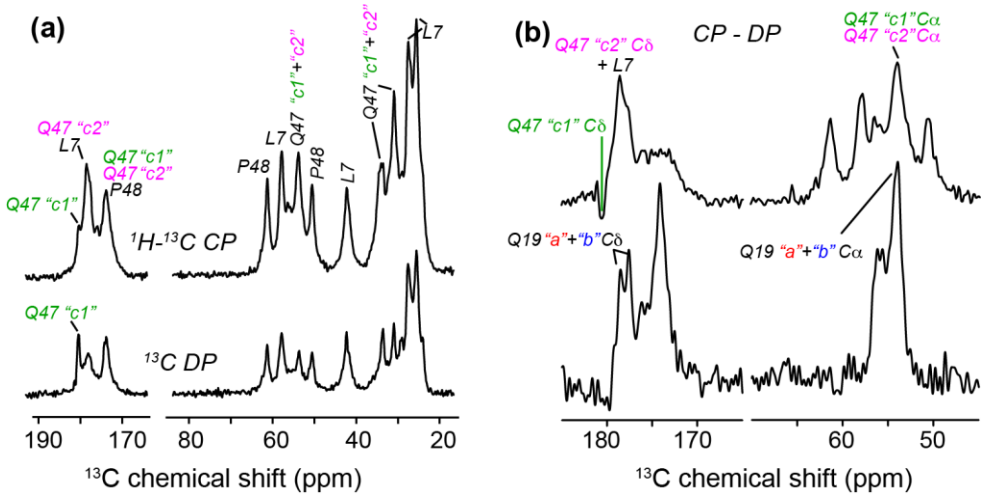

Figure 2.4 Mobility at the Q/P junction.

(a) Comparison of ^{1}H-^{13}C CP (top) and ^{13}C DP (bottom) 1D MAS ssNMR spectra on the LQP-labeled fibrils at 9.8 kHz MAS. Q47 peaks are indicated and color-coded by conformer "c1" (green) and "c2" (magenta). (b) CP-DP difference spectra for fibrils from LQP- (top) and LKSQ-labeled (13 kHz MAS) (bottom) fibrils. High intensity peaks in these difference spectra indicate increased rigidity. Q47 is less rigid than Q19 in the polyQ core, with especially pronounced mobility for the side chain carbonyl group (Cδ) of conformer "c1". Several Q47 side chain peaks, e.g. the c1 Cδ (far left in (a)) or the Cβ/Cγ signals in Fig. 3d, are also significantly narrower, indicative of fast side-chain motion.

2.3.3 Uniform $^{13}C,^{15}N$-Labeled Exon 1 Aggregates

To complement the residue-specific insights into the Gln and Pro domains enabled by use of our synthetic peptide aggregates, uniformly $^{13}C,^{15}N$-labeled htt exon 1 was prepared by over-expression in E. coli. Removal of the MBP fusion domain by factor Xa treatment released htt exon 1 with a 44-residue polyQ domain and a C-terminal His-tag. Due to a new linker design, factor Xa cleavage did not leave additional residues N-terminal to the htt^{NT} segment, as validated by ESI TOF mass spectrometry data showing the released exon 1 to have the gene-encoded start Met-Ala-Thr. Under conditions analogous to our study of the residue-specifically labeled samples, 1.1 mg of the fibrillized U-$^{13}C,^{15}N$ labeled proteins was analyzed by 2D ^{13}C-^{13}C ssNMR spectroscopy. The obtained spectrum (**Figure 2.5**) is dominated by the signals from the most rigid parts of the sample: the 44-residue polyQ domain and the many prolines in the C-terminal PRD segment. These signals match those obtained from the residue-specific labels discussed in the synthetic htt exon 1 model peptide samples. This can be readily appreciated from **Figure 2.5b**, in which spectra from the fully- and residue-specifically labeled fibrils are overlaid (see also **Figure 2.6**). Thus, it appears that the NMR signals observed for specific residues labeled in the polyQ core of the $htt^{NT}Q_{30}P_{10}K_2$ fibrils (i.e. Q18, Q19, and Q46) similarly apply to the bulk of the polyQ domain of htt exon 1. Analogously, the signals of P48 (**Figure 2.1**) match those of the U-$^{13}C,^{15}N$-labeled PRD in **Figure 2.5**. These data support our earlier analysis of the bulk natural abundance ^{13}C signals of both Gln and Pro in $htt^{NT}Q_{30}P_{10}K_2$ fibrils [21]. The htt exon 1 spectra are dominated by the large numbers of Gln and Pro residues, as may be expected. Consistent with its dynamic nature (see below), in the current size-limited sample it is difficult to detect the signals from htt^{NT}.

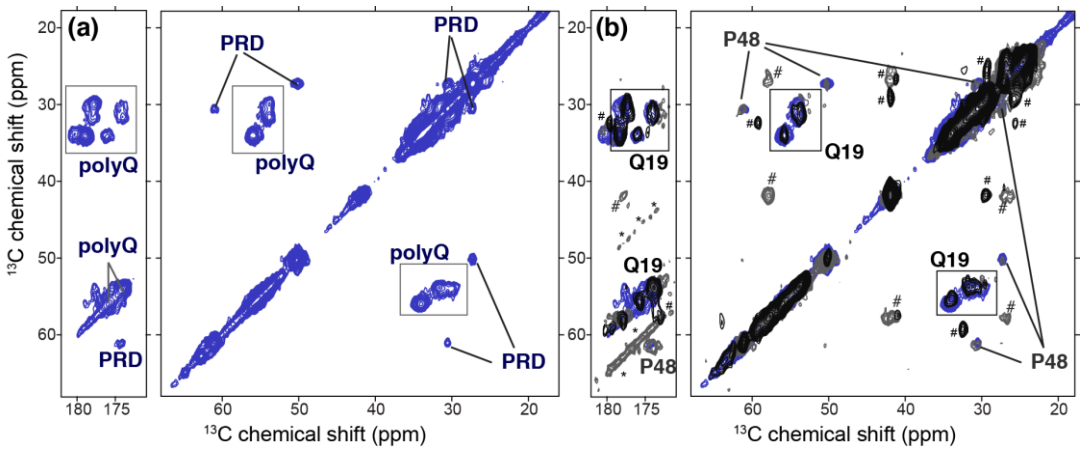

Figure 2.5 PolyQ and PRD bulk ^{13}C chemical shifts in htt exon 1.

(a) 2D $^{13}C,^{13}C$ spectrum of 1.1 mg of U-$^{13}C,^{15}N$-labeled fibrillar htt exon 1 featuring a 44-residue polyQ domain. The predominant peaks are due to the Gln in the polyQ amyloid core and the Pro residues of the PRD. (b) These peaks match those of the residue-specifically labeled htt$^{NT}Q_{30}P_{10}K_2$ fibrils from ***Figure 2.1a,d****: Q19 signals (black spectrum;* ***Figure 2.1d****) match the signal of the U-$^{13}C,^{15}N$-labeled polyQ; P48 signals (dark grey;* ***Figure 2.1a****) match those of the fully-labeled PRD prolines. Asterisks indicate spinning sidebands. Hash marks indicate htt^{NT} peaks that are visible in the htt$^{NT}Q_{30}P_{10}K_2$ sample, but not detectable in the current htt exon 1 samples (see text). All data are obtained at 600MHz, using 15 or 25ms DARR (see* ***Table A.2*** *in* ***Appendix A****).*

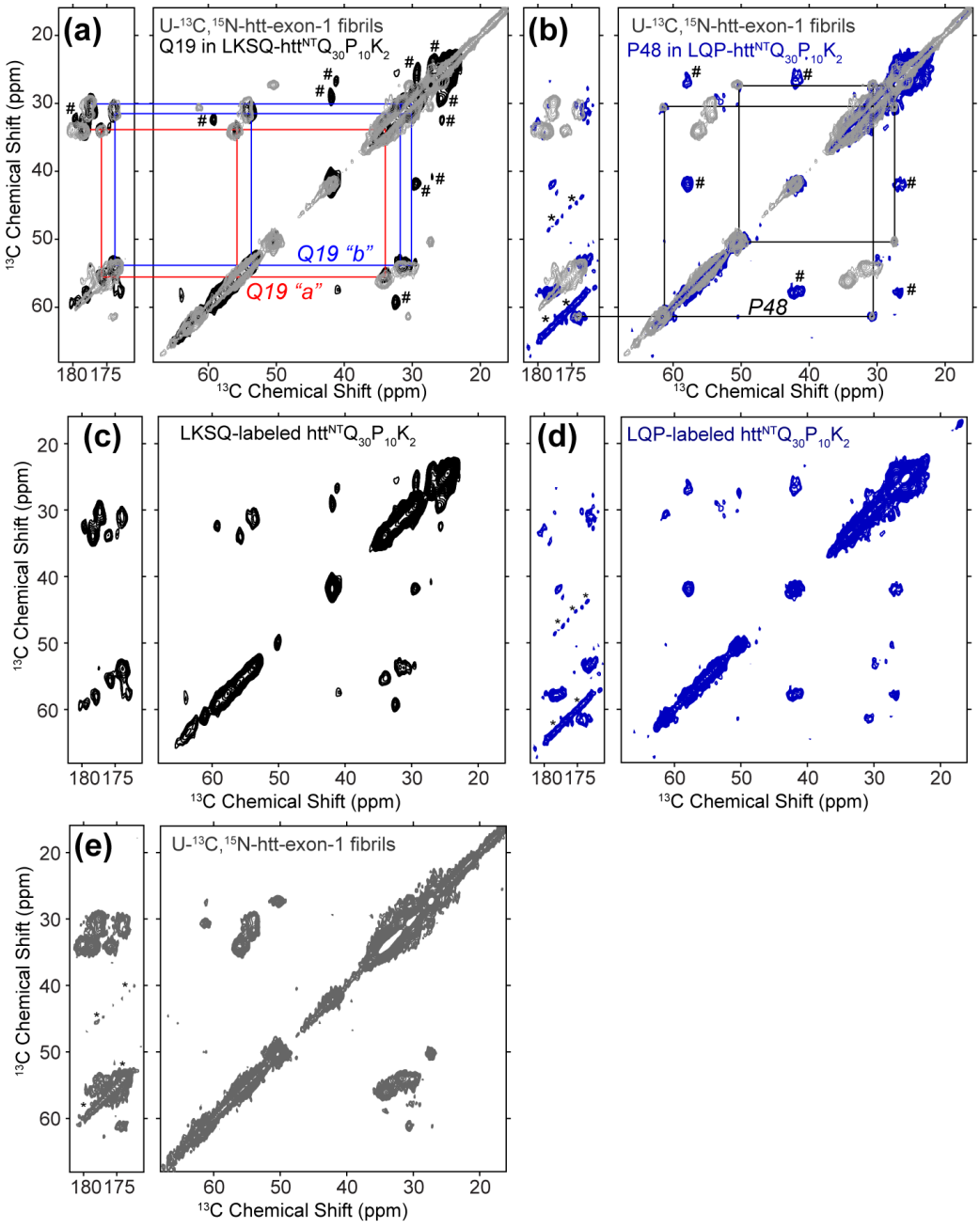

Figure 2.6 Comparison of the ^{13}C signals from polyQ and PRD in U-^{13}C,^{15}N htt exon 1 fibrils and residue-specific labels in htt$^{NT}Q_{30}P_{10}K_2$ fibrils.

(a) Overlay of htt exon 1 (gray) and LKSQ-labeled htt$^{NT}Q_{30}P_{10}K_2$ fibrils; Q19 conformers "a" and "b" are marked. (b) Overlay of htt exon 1 (gray) and LQP-labeled htt$^{NT}Q_{30}P_{10}K_2$ fibrils; P48 signals are marked. (c-d) Individual spectra for the htt$^{NT}Q_{30}P_{10}K_2$ fibrils used in the overlays in ***Figure 2.5*** *and this figure. (e) The ^{13}C-^{13}C 2D spectrum of U-^{13}C,^{15}N-labeled fibrils by itself; shown with lower contour levels close to the noise level.*

2.3.4 Dynamics of the htt^{NT} α-helix Backbone

We have previously identified the presence of a defined α-helical segment within htt^{NT} in the htt$^{NT}Q_{30}P_{10}K_2$ fibrillar aggregates, as shown in **Section 1.3.1** [21]. Those initial measurements, which for the first time demonstrated the presence of an α-helix in the fibrils and delineated its exact sequence location, were limited to qualitative measurements of dynamics due to small sample sizes. Here, we perform more detailed and quantitative ssNMR experiments on several new htt$^{NT}Q_{30}P_{10}K_2$ samples containing more peptide and new labeling patterns (**Table 2.1**). Via 2D MAS ssNMR experiments on the mature aggregates, we obtained assignments for five htt^{NT} residues: L4, K6, L7, M8, and A10 (e.g. **Figure 2.7**, **Table A.2** in **Appendix A**). The obtained assignments validated and reproduced several previous assignments [21, 54] and in all cases, the chemical shifts of the labeled htt^{NT} sites are indicative of α-helical secondary structure. Moreover, the new assignments allow for application of more quantitative torsion angle analysis (with the TALOS+ program [121]), yielding α-helical dihedral angles for L7: ϕ= -63° ± 3°, ψ= -42° ± 5° (**Figure 2.7**). These new samples also allow for quantitative analysis of the dynamics of the htt^{NT} α-helix, through the use of several complementary MAS ssNMR measurements. First we probe domain dynamics by measuring the ^{15}N longitudinal relaxation rates for the backbone nitrogens of residues in the α-helix. Based on our published [21] and current (**Figure 2.7**) ssNMR data, we know that both M8 and A10 form part of the htt^{NT} α-helix retained in the fibrils.

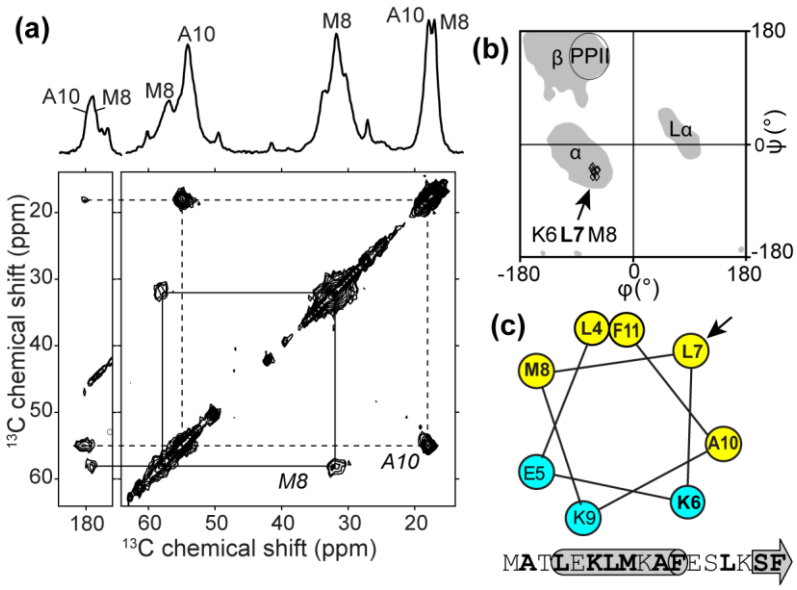

Figure 2.7 Verification of the htt^{NT} α-helix in htt^{NT}Q$_{30}$P$_{10}$K$_2$ fibrils.

(a) 1D and 2D MAS ssNMR spectra obtained on MA-labeled htt^{NT}Q$_{30}$P$_{10}$K$_2$ fibrils. Bottom: aliphatic/carbonyl (vertical section on left) and intra-aliphatic (right) regions of a 2D ^{13}C-^{13}C spectrum obtained with 8ms DARR mixing. Both spectra were acquired at 600 MHz (^{1}H freq.) and 10 kHz MAS. (b) Ramachandran plot of the backbone torsion angles for L7 (black diamonds) based on the TALOS+ analysis of the chemical shifts of K6, L7, and M8. (c) Helical wheel plot of the α-helix within htt^{NT} with ssNMR-probed residues in bold, hydrophilic residues shown in cyan and hydrophobic residues in yellow. Residue L7 (arrow) (labeled here and in ref. [21]) forms the middle of the hydrophobic face, and is thus expected to form part of a hydrophobic "core" upon clustering of the amphipathic α-helices.

Consistent with them undergoing similar (domain) motion, both residues have similar ^{15}N backbone R_1 values, i.e. 0.11 ± 0.07 s^{-1} (**Figure 2.8a**). This value exceeds the ^{15}N R_1 of structured regions of e.g. crystalline GB1 and are closer to the R_1 of solvent-exposed loop regions [122]. To test how this compares to the rigidity of the polyQ core, we also measured the backbone ^{15}N R_1 rate for site-specifically U-^{13}C,^{15}N-labeled Q10 in $K_2Q_{11}PGQ_{11}D_2$ fibrils (see also [114]). The two distinct NMR signals seen for this single amyloid core residue have

backbone ^{15}N R_1 values of 0.02-0.05 s^{-1} (**Figure 2.8b**), which resemble the structured regions in crystalline proteins.

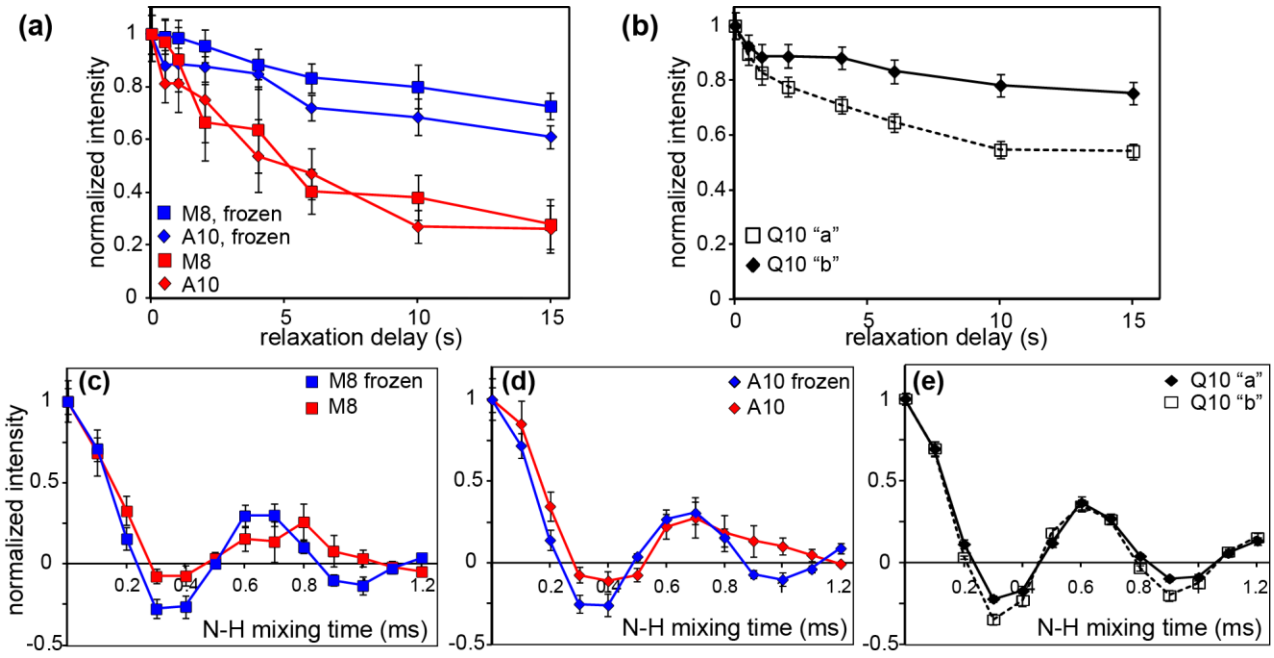

Figure 2.8 Site-specific dynamics measurements in the htt^{NT} and polyQ core.

(a,b) ^{15}N longitudinal relaxation for (a) residues in the htt^{NT} α-helix of $htt^{NT}Q_{30}P_{10}K_2$ and (b) residue Q10 of the polyQ amyloid core of $K_2Q_{11}PGQ_{11}D_2$. (a) Relaxation curves are shown for the unfrozen sample (red) and frozen sample (blue). (b) ^{15}N longitudinal relaxation curves at 275 K for conformers "a" and "b" in Q10 in unfrozen polyQ peptide $K_2Q_{11}PGQ_{11}D_2$ (described in ref. [114]). Data were obtained as 1D ^{15}N spectra at 19-22 kHz MAS. (c,d) N-H dipolar recoupling curves for M8 and A10 in the htt^{NT} of $htt^{NT}Q_{30}P_{10}K_2$ fibrils, obtained on unfrozen (287 K; red) or frozen (250 K; blue) samples, respectively. (e) N-H coupling measurement on the polyQ amyloid core, showing separate curves for both conformers of the U-^{13}C,^{15}N-labeled residue Q10 in unfrozen (275 K) fibrillar $K_2Q_{11}PGQ_{11}D_2$. (c-e) used a $R18_1^7$ symmetry sequence at 10 kHz MAS. All data were acquired at 600 MHz (^1H).

2.3.5 Order Parameters Validate htt^{NT} α-helix Dynamics

One concern with a reliance on T_1 NMR relaxation in MAS NMR as a measure of dynamics is that the T_1 relaxation may contain significant non-motional contributions, for instance due to the proximity of methyl groups that act as relaxation sinks [123-125]. This is one of the reasons why the comparison of ^{13}C CP and DP spectra that we previously applied [21] is primarily a qualitative measurement. To address this concern, we took two alternative and complementary approaches: order parameter measurements and a determination of the temperature dependence of the MAS ssNMR dynamics measurements.

NMR order parameters provide a measure of dynamics that is based on the fact that e.g. N-H dipolar couplings are reduced ("averaged") by local dynamics. We probed the backbone N-H dipolar couplings in the N-terminal α-helix using R-sequence-based DIPSHIFT experiments (see **Section 2.2.3.2**). As shown in **Figure 2.8c-e**, the time-dependent oscillations are different for residues in the htt^{NT} α-helix (**Figure 2.8c and d**) compared to a β-sheet polyQ core residue (**Figure 2.8e**). Data analysis via numerical simulations indicates significant differences in the corresponding N-H dipolar order parameters. Simulations of rigid sites (the polyQ core residue and data below) reproduced the experimental data very well and indicated dipolar coupling constants of ~11.2 ± 0.4 kHz. These are consistent with an essentially rigid state, as also seen in other amyloid-like fibrils [126]. The residues M8 and A10 in the htt^{NT} α-helix feature a decreased dipolar coupling strength (9.5 ± 1.5 kHz) and increased relaxation damping, indicative of pronounced molecular motion. For these dynamic sites, the fit quality was poorer, presumably due to an inability to fully capture the effects of anisotropic and heterogeneous dynamics. We note that the fits above did not explicitly consider the ^{1}H chemical shift anisotropy (CSA) [127].

In test simulations using standard CSA values [127, 128], the trend in coupling constants seen for the htt^{NT} and amyloid core residues was unchanged, even though the exact fit optima were dependent on the precise CSA values employed.

2.3.6 Htt^{NT} α-helix Motion is Coupled to the Solvent

To further test the connection between these NMR parameters and true dynamics, we also examined the effect of sample temperature on both the relaxation and order parameter measurements. In particular, we reduced the sample temperature until the solvent froze, as validated by observation of the broadening of the water ^{1}H NMR signals. Upon doing so, the ^{15}N longitudinal relaxation of both labeled htt^{NT} α-helix residues slowed down to R_1 values of 0.02 ± 0.03 (for M8) and 0.03 ± 0.04 s^{-1} (A10), which match those of residues in a structured polyQ core (**Figure 2.8a**, blue). In addition, a dramatic change in the N-H dipolar recoupling experiments is observed (**Figure 2.8c and d**, blue). The low-temperature N-H dipolar coupling data feature larger dipolar coupling constants (~10.5 ± 0.3 kHz) and reduced relaxation damping, resulting in curves that more closely resemble the rigid polyQ core (compare to **Figure 2.8e**).

The effects of freezing the sample are also manifest in ^{13}C-based ssNMR experiments, where we observe an increased ^{1}H-^{13}C CP intensity and a larger CP-DP difference in backbone and aromatic sites (**Figure 2.9**). This indicates rigidification of the backbones and side chains and also shows that these difference experiments are indeed reflective of increased dynamics. In other words, previously noted differences between the polyQ amyloid core and htt^{NT} segment [21] are not simply due to absence and presence of methyl groups in either domain.

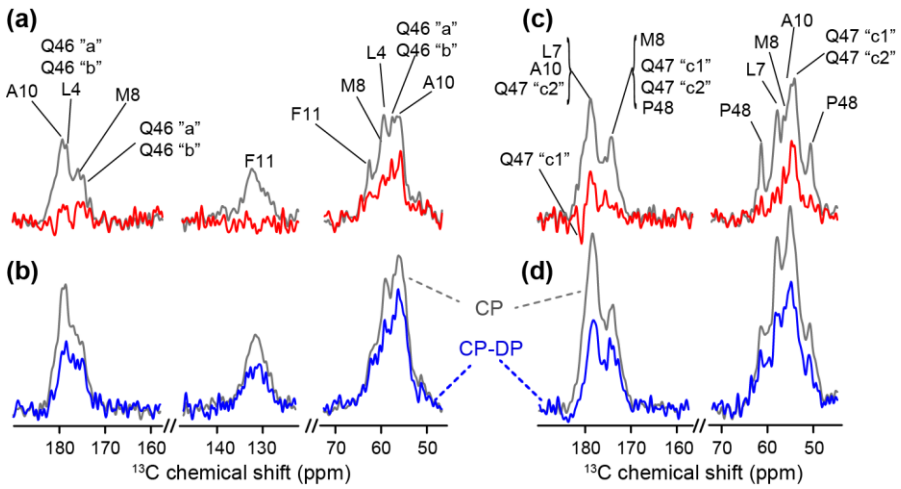

Figure 2.9 Solvent-Couple Dynamics in the htt^{NT}.

1D ^{13}C ssNMR on fibrils of co-mixed (1:1) MF-labeled/LAQ-labeled peptides at (a) 287 K and (b) 265 K and mixed MA-labeled/LQP-labeled peptides at (c) at 287 K and (d) 270 K. A CP-DP difference spectrum is shown in red/blue, with the corresponding ^{1}H-^{13}C CP spectrum in gray. The relative increase in intensity in the difference spectra, when the solvent freezes, indicates a rigidification that applies to the backbone as well as (hydrophobic) side chains (e.g. F11 in panel b). This implies solvent-coupled dynamics that extend to the hydrophobic face of the amphipathic α-helix (ref. ***Figure 2.7****).*

2.4 DISCUSSION

2.4.1 Secondary Structure of Fibrillar PolyQ and Its Flanking Domains

In monomeric polypeptides, the oligoPro flanking segment adopts a PPII conformation that modulates the structure of multiple preceding Gln residues [65, 108-111], presumably enabled in part by Gln's innate propensity for PPII structure [111, 129] (**Figure 2.10**). Our MAS ssNMR data (**Figure 2.1** and **Figure 2.2**) show that the oligoPro segment in our $htt^{NT}Q_{30}P_{10}K_2$ fibrils also adopts a PPII helix, which starts from the very first Pro residue. The same resonances are also seen in **Figure 2.5** and **Figure 2.6** for Pro residues in U-^{13}C,^{15}N-labeled htt exon 1 fibrils. We also found that the Gln residue preceding the PRD fails to adopt the canonical β-sheet polyQ amyloid conformation. This is indicated by the complete absence of the characteristic polyQ ssNMR signature (**Figure 2.1**), as well as increased motion and solvent accessibility of this residue (**Figure 2.3**). Remarkably, unlike earlier solution-state studies, we find that in the fibrils this oligoPro-induced non-β-structure is mostly limited to a single Gln. In the fibrils, the penultimate Gln, Q46, does display the chemical shift signature of the polyQ amyloid core. A minor population (~25%) of Q46 is found to show some of the characteristic side chain chemical shifts that are seen for Q47 and attributed to mobile residues outside the amyloid core [22] (see also below). In the U-^{13}C,^{15}N-labeled htt exon 1 fibrils we find that the bulk of the polyQ domain features the same two sets of ssNMR signals. Thus, the amyloid core is formed by Gln residues throughout most of the polyQ segment. These data are in line with our previous analysis [21] of the bulk Gln signals in unlabeled $htt^{NT}Q_{30}P_{10}K_2$ fibrils. They are also reminiscent of a ssNMR

study of aggregated uniformly labeled polyQ without htt flanking domains [22]. Thus, once formation of amyloid structure is initiated (either by spontaneous nucleation or via seeded elongation by pre-formed polyQ amyloid), β-sheet structure propagates throughout essentially the entire polyQ domain. As may be expected, there is no indication of β-sheet structure in the oligoPro segment, in contrast to the formation of some β-structure in the final residues of htt^{NT} (see [21]). Thus, the "protective" ability of the oligoPro segment to induce PPII structure within multiple preceding Gln is specific to the monomeric (and possibly oligomeric) species, and is not retained in the mature amyloid aggregates.

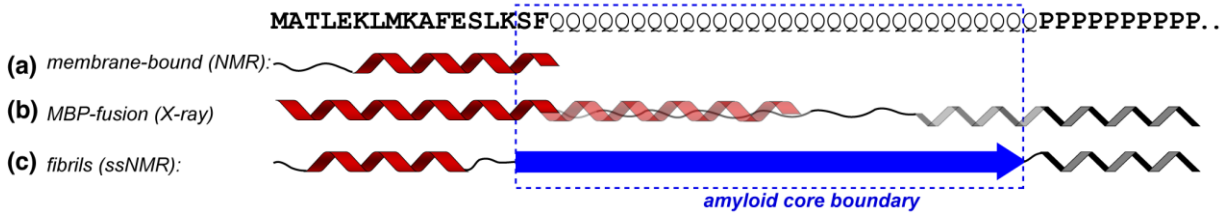

Figure 2.10 Secondary structure of htt N-terminal fragments under differing conditions.

*(a) NMR studies have shown extensive α-helicity in isolated htt^{NT} upon interacting with a lipid bilayer[130]. (b) In crystals formed from a MBP-fusion construct, htt^{NT} is fully α-helical, followed by partial α-helicity in the N-terminal part of the polyQ domain [131]. The oligoPro segment forms a PPII helix, with extended or PPII structure for several preceding Gln residues. (c) Mature fibrils feature a relatively well-defined β-sheet amyloid core (dashed blue box; see also **Figure 2.11**), formed at the expense of any α-helical or PPII structure within the polyQ. A short α-helix in htt^{NT} helps stabilize the fibrils via molten-globule-like assemblies that may be inherited from non-β oligomers.*

By ssNMR, we find that in fibrils an α-helix in htt^{NT} spans only part of this N-terminal segment (**Figure 2.10**, **Figure 2.7**; ref. [21, 54]) and undergoes extensive dynamics (e.g. **Figure 2.8**). Although various studies have proposed or discussed α-helical structure for htt^{NT} [47, 131-134], these studies did not take place in context of fibrils. An intriguing feature of the fibrillar

htt^{NT} α-helix is that it is shorter than α-helices seen in experimental studies in the context of non-oligomers, membranes, and the crystal lattice of a chimeric construct (**Figure 2.10**) [54, 130, 131, 135]. The detailed conformation of the oligomers remains uncertain, but it is clear that interactions between htt^{NT} α-helices are critical for their formation [21, 49, 54, 91, 93, 135]. At the same time, intermolecular interactions seem to be required to have htt^{NT} form stable α-helical structure [49, 91, 135].

Previous studies have reported that the polyQ domain may have a propensity for α-helical content [131, 133, 134], especially in the N-terminal polyQ portion when attached to an α-helical flanking domain [65, 131, 136, 137]. As one example (**Figure 2.10b**), extensive α-helical structure is seen within the polyQ domain of N-terminal htt fragments fused to MBP [83]. Once fibril formation has completed, our use of MAS ssNMR on fibrils shows that no α-helical structure remains in the polyQ domain [21, 54]. Indeed, this β-structure actually extends into the C-terminal end of the htt^{NT} segment, at its C-terminal junction with the polyQ (**Figure 2.10**).

2.4.2 Boundaries and Alignment of the PolyQ Amyloid Core

We previously described [114] how the β-sheet polyQ amyloid core can be identified based on its unique chemical shift signature featuring the two conformers "a" and "b" (as shown in **Section 1.3.2**. and illustrated in **Figure 2.1** in red and blue, respectively). Through our use of synthetic peptides and residue-specific labeling, we have shown that these two conformers are observed even for a single labeled residue independent of its position in the polyQ domain. In htt^{NT}Q$_{30}$P$_{10}$K$_2$ aggregates, this signature was seen when we labeled the first two Gln residues directly following the htt^{NT} flanking domain [21]. Here, we show it is also present near the C-

terminal end of the polyQ domain, providing further support to our hypothesis that this doubled signature is a general and intrinsic feature of Gln throughout the β-strands of polyQ amyloid. Moreover, we also showed that the same chemical shifts are seen for the Gln residues that make up the bulk of the polyQ amyloid core in the U-^{13}C,^{15}N labeled htt exon 1 fibrils (**Figure 2.5**, **Figure 2.6**). This complements and confirms our previous analysis of unlabeled htt$^{NT}Q_{30}P_{10}K_2$ fibrils, based on spectroscopic filtering of the rigid amyloid core [21]. Work by the Baldus group on polyQ fibrils lacking htt-like flanking domains reproduced our assignments of the amyloid core signals, even though, due to their labeling approach, it was not apparent that the doubled signature is an attribute of each individual amyloid core residue [22]. Further ssNMR studies into the conformational underpinnings of this polyQ-specific amyloid [25] signature are presented in **Chapter 3.0** , which we here employ to establish the presence or absence of specific residues within the amyloid core.

In many amyloids, one finds well-defined boundaries for the fibrils' β-sheet amyloid core. In these amyloids, β-strand-based "unit cells" within the β-sheets stack into a single well-defined register that balances the alignment of hydrophobic residues, the formation of Asn or Gln "ladders" and the interactions of charged residues. This is perhaps most obvious in the commonly found parallel in-register amyloid structures, where the β-sheets are defined by translation of a unit cell featuring a single β-strand [138-140]. Also more complex cases, such as the fibrils formed by the fungal prion HET-s [141] or the antiparallel fibrils made by Iowa-mutant Aβ [142] feature a strictly aligned repeating unit within the fibrillar β-sheets (see **Figure 2.11a and b**). Misalignment of new β-strands that are added to the elongating fibril has a significant energetic penalty, resulting in the templated elongation into structurally homogeneous fibrils. However, during elongation of polyQ amyloid, the alignment constraints on incoming

polypeptide chains are much less rigorous due to the repetitive and degenerate nature of the primary sequence. This translates into an ability to incorporate polyQ domains of other lengths and from other proteins – a phenomenon previously described as "promiscuous" polyQ amyloid growth [23, 24, 143-145]. Even when a homogeneous pool of polyQ peptides is forming amyloid, the penalty of having incoming peptide chains not lining up perfectly with the existing amyloid core is likely to be relatively low, especially for longer polyQ lengths. Thus, this would allow for a "misalignment" (from the point of view of residue number, if not of residue type) of the incoming peptide chains with the existing polyQ amyloid core. This translates into having peptide chains in various residue number "registers" inside each individual fibril. Among other effects, this kind of chain misalignment could lead to the generation of disordered loops and tails decorating the amyloid core, with lengths and locations that vary from peptide to peptide, even within one fibril. A previous in silico study has noted this kind of imperfect alignment and how it resulted in disordered tails at the ends of the polyQ peptide [86]. In ssNMR experiments, such disordered non-β segments would manifest themselves as Gln signals lacking the amyloid core signature and featuring enhanced dynamics. Such signals were observed in a previous ssNMR study [22] on polyQ fibrils lacking htt-based flanking domains (see also below). Although these were discussed in context of flexible loops, they would also be expected for "tail ends" of disordered Gln beyond the polyQ core. Here, we find in htt context that the U-^{13}C,^{15}N-labeled fibrils also contain this type of signal, with it accounting for 11% of the Gln based on an integration of the Cδ-Cγ/Cβ cross-peaks. In the site-specifically labeled htt$^{NT}Q_{30}P_{10}K_2$ fibrils we observe that the final Gln (Q47) has chemical shifts that closely match those of the disordered glutamines. On the other hand, for the very first Gln residue of the polyQ segment (Q18) [21] and also the penultimate Gln (Q46), both of which form part of the amyloid core, we find only

minor populations of the disordered Gln residues (~15% and 25%, respectively). The second Gln residue, Q19, lacks such a disordered population (**Figure 2.1i**). These observations seem consistent with a very limited degree of residue number misalignment within the polyQ amyloid core. It remains unclear whether this is the full explanation of these observations. Generally, we note that the proportion of disordered glutamines in these fibril samples seems to be lower than that reported for polyQ fibrils lacking htt flanking domains [22]. This might imply an ordering effect of the non-β htt^{NT} and oligoPro flanking sequences, which could effectively help "set the register" of the polyQ amyloid core [82]. Alternatively, this may also in part reflect a dependence on the polyQ length, since in the abovementioned study the disordered Gln pool seemed to be most pronounced in the longest polyQ [22]. Moreover, the degree of structural and dynamic disorder within the fibrils' polyQ domains can be affected by the aggregation conditions, as previously discussed for htt exon 1 aggregates [62].

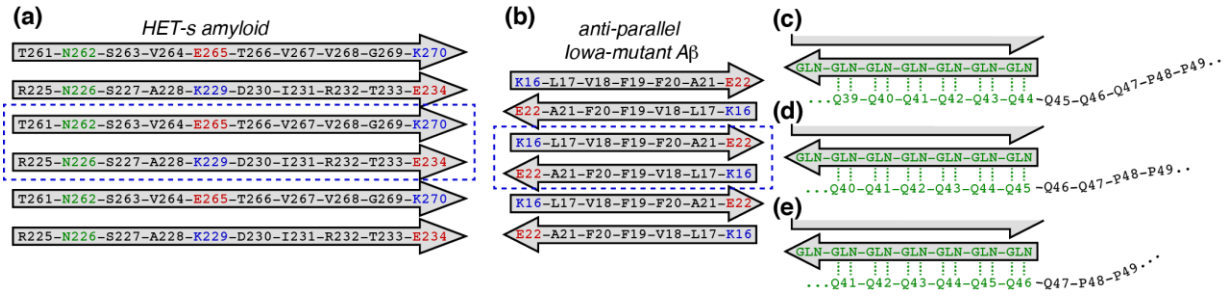

Figure 2.11 Peptide chain (mis)alignment during amyloid elongation.

The amyloid cores of (a) HET-s and (b) Iowa mutant Aβ are assembled from strictly aligned repeating "unit cells" (dashed rectangles) [141, 142]. Incoming peptide chains align their charged (red/blue) and hydrogen-bonding side chains (Asn/Gln; green). (c-e) The anti-parallel β-sheet polyQ amyloid is more forgiving of misalignment of incoming peptide chains, since a different alignment with the amyloid core (green) still allows for similar backbone- and side-chain hydrogen bonding to the core (dashed lines). This would leave varying numbers of non-amyloid Gln at the N- and C-terminal polyQ ends (as shown in (c-e)) and in flexible loops. This may explain the localized

*disorder in Gln at or near the polyQ end (**Figure 2.1i**). These non-amyloid signals are largely (for 75% of the fibrillized peptides) limited to the final Gln, consistent with the schematic shown in (e).*

The mobile, non-amyloid-core Gln residues that were identified as the "Q(3)" population by the Baldus group [22] feature a distinct chemical shift pattern. It can be distinguished from the polyQ core residues, by its Cδ chemical shift near 180 ppm (**Figure 2.1e**; green bars at the far right), or the presence of an off-diagonal Cβ-Cγ peak in ^{13}C-^{13}C 2D spectra (**Figure 2.1b**). A close inspection of several other Gln-labeled fibrils that we have studied (refs. [21, 54, 114] and unpublished results) reveals weak signals with similar chemical shifts. This includes both the U-^{13}C,^{15}N exon 1 fibrils' bulk Gln signals and specifically to the first and last Gln residues in our htt^{NT}Q$_{30}$P$_{10}$K$_2$ fibrils. Perhaps most strikingly, this applies to residue Q47, which has peaks (in particular conformer c1) that closely resemble these chemical shift values (**Figure 2.1e**). The most notable deviation is in the ^{13}CO chemical shift, but this is likely due to the well-established effect of the adjacent Pro [146]. The Q47 chemical shifts also bear strong resemblance to those of Gln preceding Pro in intrinsically disordered proteins (IDPs), which have a characteristic set of backbone chemical shifts of ~174 ppm (CO) and ~53.5 ppm (Cα). These differ significantly from both random coil shifts and from non-pre-proline IDP Glns. Therefore, the final Gln is indeed disordered and affected by the PPII oligoPro segment. We speculate that Q47 might feature a PPII-like structure that resembles the conformation of the oligoPro-flanked polyQ tail prior to β-nucleation.

2.4.3 Htt-Binding Proteins and Their Effect on Aggregation

Thus, we localized at the Q/P junction a transition from β-sheet to PPII structure, occurring remarkably abruptly at the final Gln residue, which lacks the amyloid core signal and instead features increased dynamics and water exposure. The PPII-structured oligoPro elements are known to provide the binding sites for certain htt-binding proteins, some of which are able to bind to fibrils [116]. One of these, the PPII-binding WW domain protein FE65, is thought to specifically recognize this particular Q/P junction, which would require the observed accessibility (i.e. solvent exposure) of said junction [118]. The anti-oligoproline antibody MW7 can bind aggregates from mice [117, 147-149], and it can also destabilize pre-existing fibrils and even interfere with fibril formation [149]. These experiments indicate that one of its oligoPro binding sites occurs at the very edge of the amyloid core. Optimal tight binding of MW7 to this site may thus interfere with the formation of the polyQ amyloid core or at least reduce the stability of the fibrillar assembly.

2.4.4 Molten-Globule-Like Dynamics of the α-helical htt^{NT} Segment

Intermolecular htt^{NT} interactions play a crucial role in accelerating the aggregation of htt fibrils, by enabling the formation of α-helix-rich oligomers [48, 49, 91, 135]. Htt^{NT}-htt^{NT} interactions also contribute to the stability of the mature fibrils [23, 54]. Our MAS ssNMR data show that in mature fibrils the α-helical segments nonetheless remain at least partly solvent-exposed. Moreover, despite the structural role in the fibrils [23, 54], the htt^{NT} α-helix experiences solvent-coupled dynamics, both in its backbone and hydrophobic side chains. This is reflected in greater ^{15}N and ^{13}C longitudinal relaxation rates and reduction of N-H dipolar

couplings. The pronounced effect of freezing the solvent shows that this mobility is coupled closely to the solvent dynamics. Solvent freezing even impacts the hydrophobic face of the amphipathic helix (e.g. Phe ring in **Figure 2.9**), which would form the hydrophobic core of clustering of the amphipathic α-helices within the fibrils. Thus, our ssNMR data indicate that the interacting htt^{NT} α-helices in the fibrils are in a molten-globule-like state, in which amphipathic α-helices are formed and interact but remain flexible and loosely packed. Although we still lack a good understanding of the structure and dynamics of htt-related oligomers, such molten-globule-like htt^{NT} assemblies remind us of the proposed molten-globule-like intermediate states of other amyloid-forming proteins [150-154]. If the α-helix-rich oligomers formed by htt N-terminal fragments turn out to be dynamic, then the dynamic htt^{NT} α-helices in the fibrils might even trace their origins to the initial interactions in the self-assembly of the oligomeric precursors.

2.5 CONCLUSIONS

Extensive structural changes are common during the aggregation of amyloidogenic proteins. This includes the loss of native secondary and tertiary structure elements as well as the formation of the highly rigid, intermolecularly hydrogen-bonded β-sheets that characterize amyloid-like fibrils. Through the use of MAS ssNMR spectroscopy, we have gained new insights into the structural features of the polyQ domain and its flanking segments after the aggregation of htt N-terminal fragments. Following nucleation, a β-sheet-based amyloid core is formed in the polyQ domain, which includes Gln residues very close to the C-terminal oligoPro segment. The latter forms a PPII helix that extends to, but fails to significantly disrupt, the polyQ amyloid core in mature fibrils. This contrasts with its effect prior to fibril formation, when it modulates the conformational ensemble of multiple preceding glutamines sufficiently to suppress nucleation. The β-sheet amyloid core extends slightly beyond the N-terminal end of the polyQ segment, stopping short of the amphipathic α-helix close to the very N-terminus [21]. Here, we found that this amphipathic helix displays pronounced solvent-coupled dynamics, seemingly reflecting a molten globule-like assembly that is reminiscent of some oligomeric amyloid precursors. Whilst α-helix-rich and dynamic oligomeric intermediates are implicated in other amyloid-forming proteins and peptides [153-156], htt-related fibrils may be relatively unusual for such proteins in that they retain a substantial part of these structural and dynamic features in the mature fibrils. The obtained new insights into the structure of htt-related fibrils may also help explain reported inhibitory effects and aggregate binding characteristics of antibodies and other htt-binding proteins.

2.6 ACKNOWLEDGEMENTS

We thank Mike Delk, Guy Uechi, Ravindra Kodali and Bankanidhi Sahoo for technical assistance and helpful discussions.

3.0 CONFORMATIONAL UNDERPINNINGS OF THE POLYGLUTAMINE SPECIFIC AMYLOID SIGNATURE

Hoop, C.L., Lin, H.K., Karunakar, K., Lamley, J.M., Lewandowski, J.R., Wetzel, R., and van der Wel, P.C.A. (2014) Conformational underpinnings of the polyglutamine specific amyloid signature, *manuscript in preparation*.

Author contributions: C.L.H. acquired and analyzed MAS ssNMR data on all samples. H.K.L. expressed, purified, and fibrillized [U-^{13}C,^{15}N]-htt exon 1 and acquired MAS ssNMR spectra. K.K. purified and fibrillized all synthetic peptides. J.M.L. and J.R.L. acquired and analyzed ultrafast spinning relaxation measurements. C.L.H. and P.v.d.W. wrote the manuscript.

Huntington's disease is one of at least ten neurodegenerative diseases caused by a genetic mutation leading to an expansion of CAG repeats, which encodes for polyglutamine (polyQ) tracts that self-aggregate to form fibrils. Molecular details of the fibril assembly of the polyQ core are critical to understand the aggregation pathway. Thus far, in various contexts, polyQ amyloid core Gln residues have shown a characteristic signature motif NMR chemical shift pattern that seems unique to polyQ. These Gln residues show doubled chemical shifts representing equal distributions of two major conformers. Here, we explore these conformers in mature fibrils of polyQ-containing peptides and full-length htt exon 1. We use magic angle spinning (MAS) solid state NMR (ssNMR) to determine structural and dynamical details of these

two Gln conformers, in order to detail their connectivity, constrain backbone and side chain torsion angles, and measuring ^{15}N and ^{13}C relaxation. These parameters allow us to provide a new understanding of the structure of the polyQ amyloid core. We also examine how incorporation of L-Pro-Gly and D-Pro-Gly motifs inserted into polyQ domains affect the local structure and facilitate the formation of β-hairpin structures. These observations together provide new insight into the possible misfolding and aggregation pathway of expanded polyQ, while providing a molecular explanation for its unique doubled ssNMR chemical shift signature.

3.1 INTRODUCTION

At least ten human disorders are caused by a genetic mutation involving the expansion of CAG repeats, which encode for polyglutamine (polyQ) tracts in associated proteins [44, 57, 58]. In such diseases, expansion of the number of glutamines beyond a threshold correlates to a decrease in the age of pathological onset and an increase in toxicity [58]. Huntington's disease (HD) is the most common of these diseases, with the Gln expansion occurring in exon 1 of the huntingtin protein (htt). The mutant proteins in HD and other polyQ expansion disorders undergo a misfolding and self-assembly processes, resulting in non-native conformations that cause cytotoxic effects. As in other protein misfolding disorders, modulation of the misfolding process might thus be a beneficial treatment or preventative strategy. The effective design of targeted drugs is hindered by the lack of unequivocal and detailed data on the molecular events that characterize misfolding, oligomer formation and fibrillization.

Understanding the structure of the misfolded aggregates implicated in HD is an important

and necessary step toward characterizing the aggregation pathway. The amyloid-like aggregates observed in HD contain N-terminal fragments of the htt protein. Among these, a fragment that represents the protein's exon 1 segment has been shown to induce HD-like disease pathology in mouse models. This has resulted in it becoming a commonly used reference point in functional, molecular and biophysical studies. Structural studies of polyQ aggregates have provided a number of insights into the internal structure of the assemblies, but a detailed molecular structure is still missing. Unlike most Aβ fibril forms, polyQ fibrils do not arrange in typical in-register, parallel β-sheets but feature instead anti-parrallel β-sheets [21, 23]. Several different molecular models have been introduced, based on diverse experimental approaches. However, these models lack atomic detail and do not agree with each other. Proposed models include water-filled nanotubes [26], fibrils containing β-hairpins with extended [19] or bent side chains[20], and super-pleated β-arches [22]. These models vary in the global architecture in which the β-strands are arranged. They also differ on a local atomic level, as reflected in the Gln side chain torsion angles within the assembly. This is of note, since for many spectroscopic techniques, the observed resonance frequencies are ultimately determined by these dihedral angles. Thus, the models differ both in the overall assembly and on the atomic level, reflecting the paucity of concrete experimental constraints on the fibrillar structure.

Magic-angle-spinning (MAS) NMR spectroscopy is a rapidly developing technique that is capable of atomic-level structural and dynamical measurements on non-crystalline assemblies, such as amyloid-like fibrils. It has been used to determine the molecular structure of amyloid fibrils [5, 45, 157, 158]. Due to their dependence on the atomic structure, MAS NMR spectra also offer a unique tool to detect and characterize the structural polymorphism of amyloid fibrils. Previously, MAS NMR studies have probed the Gln residues in various aggregated polyQ-based

peptides or proteins [21, 25, 45, 54]. Remarkably, similar NMR signals are found for polyQ amyloid core residues in polyQ constructs of different lengths, in the context of synthetic htt N-terminal fragments, and htt exon 1[21, 22, 25, 45]. When in the rigid amyloid core, each individual Gln has two distinct sets of chemical shifts. We have previously reported that the chemical shift pattern is seemingly unique, observed in <2% of all Gln residues in protein chemical shifts reported to the Biological Magnetic Resonance Bank [25].

Determination of the molecular structure of the polyQ-based fibrils may allow for an improved understanding of the molecular underpinnings of their observed toxicity to neuronal cells. One proposed contributing mechanism is based on recruitment of essential non-mutant proteins, via a seeding process that is dependent on the polyQ aggregate structure [59]. In addition, a better molecular understanding of the aggregate structure may shed light on preceding molecular rearrangements and in particular the initial nucleation events that trigger the aggregation process. An interesting aspect of simple polyQ peptide aggregation is that by most accounts, the aggregation process proceeds without the formation of oligomeric intermediates. Thus, it seems possible that key features of the initial nucleated species are retained in the mature fibrils. Modulation of the early stages of aggregation process are key targets for drug development.

Various studies have focused on the formation of β-hairpin-based structures during the misfolding of the expanded polyQ [159, 160]. The idea that β-hairpin structures form in monomeric polyQ goes back to the early work of Perutz and co-workers [161]. Later work, by the Wetzel group and others, introduced Pro-Gly insertions into polyQ segments in order to probe polyQ misfolding, aggregation as well as toxicity [24, 137, 162-166]. These kinds of motifs are expected to disrupt β-strands [137, 162], while favoring the formation of β-turn

structures. In line with a role for β-hairpins in the aggregation process, Pro-Gly and other β-hairpin stabilizing mutations preserve or enhance aggregation kinetics while also enhancing fibril stability [25, 167]. Moreover, both Pro-Gly-interrupted polyQ peptides and htt exon 1 retain cytotoxicity. Looking at different β-hairpin stabilizing mutations, the effects on aggregation kinetics and fibril stability both correlate with the degree of β-turn stabilization. This is for instance apparent in a comparison between L-Pro-Gly and D-Pro-Gly insertions. The latter has a much stronger effect, in line with its enhanced ability for β-hairpin stabilization [24, 25].

A systematic study incorporating Pro-Gly into polyQ peptides found that the amyloid core could accommodate a multitude of such insertions. As long as the motifs were nine or more Gln apart, they did not slow aggregation kinetics, even upon inclusion of three such motifs in a 46- residue long PGQ_9 construct [24]. Subsequent studies have shown that similar constructs are accommodated in an htt context, have cellular toxicity, and feature similar aggregation kinetics and aggregate morphology [137, 163-166]. Although the Pro-Gly insertions are generally assumed to form β-turns in the polyQ context, this assumption has not been universal. It has recently been suggested that Pro insertions may be accommodated in the β-strand structure, rather than the widely assumed β-turn [22]. This is not unprecedented, as MAS NMR studies of a transthyretin fragment have shown a Pro residue to be incorporated within the fibrils' β-strands. Nonetheless, such a model seems hard to reconcile with the various experimental data, in particular the observed correlations between β-hairpin stabilizing mutations and the stability of the resulting fibrils [25].

One of the intriguing consequences of the introduction of β-hairpin stabilizing mutations reflects their impact on the nucleation event. Kinetic studies have shown that short polyQ peptides feature multimeric nuclei, whereas expanded polyQ aggregates via a monomeric

nucleus. Incorporation of a β-hairpin enhancer leads to a decrease in the nucleus size down to a monomeric nucleus even for shorter polyQ peptides [25].

Here, we use a MAS NMR spectroscopy to test the universality of the doubled polyQ ssNMR signature. First we examine $K_2Q_{15}K_2$ fibrils, which aggregate via a multimeric nucleus and have been argued to not show the peak doubling. We also extend our studies of htt exon 1 fibrils and peptides featuring both D-Pro-Gly and L-Pro-Gly insertions. These experiments further support the universal nature of the doubled polyQ signature. Through backbone and side-chain torsion angle measurements we probe the structural underpinnings of the signature. Our NMR data argue against hypothesized α-sheet structures in the polyQ amyloid core [168]. Instead, we find that the amyloid core is assembled from two types of β-strands that have distinct dynamics but both feature extended side chain conformations. Finally, we show that the D- and L-Pro-Gly insertions have quite different impacts on the local polyQ structure, but preserve the amyloid core structure. We integrate these results to provide a molecular rationale for the observed polyQ ssNMR signature, which is informed by mechanistic and kinetic experimental data.

3.2 EXPERIMENTAL PROCEDURES

3.2.1 Preparation of PolyQ Peptide Fibrils

To allow for residue-specific isotopic labeling, several polyQ-based peptides were prepared by solid-phase peptide synthesis (**Table 3.1**). Peptides were obtained crude and purified in-house using established and previously reported methods [25, 95]. Fmoc- and side-chain-protected ^{13}C, ^{15}N-labeled amino acids were purchased from Cambridge Isotope Laboratories (Andover, MA) and Isotec (Sigma-Aldrich, St. Louis, MO). $D_2Q_{15}K_2$ peptide (*Q15*) was obtained from Anaspec (Fremont, CA), featuring a single U-^{13}C,^{15}N-labeled Gln residue in position 6 (i.e., the fourth Gln) and an acetylated N-terminus. The remaining peptides were synthesized at the Small Scale Synthesis facility at the Keck Biotechnology Resource Laboratory of Yale University. Several peptides were obtained featuring previously described Pro-Gly-based insertions [25] into a 22-residue polyQ domain. Peptides designated *pG1* and *pG2* featured a sequence $K_2Q_{11}pGQ_{11}K_2$ with a central D-Pro-Gly motif ('p' = D-Pro). Peptide *pG1* had two U-^{13}C,^{15}N-labeled Gln residues in positions Q8 and Q9, away from the site of D-Pro-Gly insertion. Peptide *pG2* instead featured its U-^{13}C,^{15}N-labeled residues around the site of mutation (residues Q13, G15, and Q16). Peptides designated *PG1* and *PG2* with sequence $K_2Q_{11}PGQ_{11}D_2$, instead featured a central L-Pro-Gly insertion. These peptides were also U-^{13}C,^{15}N labeled around and within the mutation site, in Q13, P14 and G15 (*PG1*), or P14 and G15 only (*PG2*). Prior to fibril formation in PBS buffer (pH 7.4) at 37°C, all peptides were disaggregated as previously described [25, 95]. Aggregation was monitored using an HPLC-based sedimentation assay and

allowed to proceed to completion [25, 95]. Uniformly ^{13}C, ^{15}N- labeled htt exon1 fibrils were prepared by Dr. Hsiang-Kai Lin as stated in **Section 2.2.2**.

Table 3.1 Sequences, labeling, and amounts of isotopically labeled NMR samples

ID	Labeling	Sequence [a)]	Sample size (mg)
Q15	U-^{13}C, ^{15}N-[Q6]	$D_2QQQQQ_{11}K_2$	5
PG1	U-^{13}C,^{15}N-[Q13, P14, G15]	$K_2Q_{10}QPGQQ_{10}D_2$	10
PG2	U-^{13}C,^{15}N-[P14, G15]	$K_2Q_{10}QPGQQ_{10}D_2$	7
pG1	U-^{13}C, ^{15}N-[Q8, Q9]	$K_2QQQQQQQQpGQQ_{10}K_2$	9
pG2	U-^{13}C,^{15}N-[Q13, G15, Q16]	$K_2QQQQQQQQpGQQ_{10}K_2$	7

3.2.2 MAS Solid-State NMR

For most samples, 5-10 mg of mature fibrils (**Table 3.1**) were pelleted into 3.2 mm zirconia MAS rotors (Bruker Biospin, Billerica, MA and CortecNet, Voisins-le-Bretonneux, France). Prior to packing, the polyQ peptide fibrils were washed with deionized water. Sample packing made use of a home-built ultracentrifugal sample packing tool, operated typically at up to 130,000 x g in a Beckman Coulter Optima L-100 XP ultracentrifuge with a SW-32 Ti rotor. Samples for the ultrafast (60k Hz) MAS NMR experiments were packed into 1.3 mm MAS rotors (Bruker). To maintain sample hydration, the 1.3 mm rotors were sealed using silicone-based spacers, leaving a total sample volume of ~1 μL, including ~0.5-1 mg of fibrils. The sample used for the carbon $R_{1\rho}$ and ^{13}C R_1 relaxation measurements was packed into a 1.3 mm rotor using a centrifuge. For the less sensitive ^{15}N measurements, a separate sample was packed with home-built tools (based on those described in [169] and [170]) in a Beckmann L8-70M ultracentrifuge running at ~100,000 g for 6 hours. This permitted an improvement of signal to

noise by a factor of ~2. Before, during and after packing, the fibril samples were kept hydrated and unfrozen at all times.

3.2.2.1 MAS ssNMR Assignment Experiments

Most MAS ssNMR experiments were performed with a widebore Bruker Avance I NMR spectrometer operating at 600 MHz ^{1}H Larmor frequency (14.1 T) and a 3.2 mm Efree HCN (Bruker Biospin) probe, unless stated otherwise. The sample temperature was controlled using a constant flow of cooled gas of ~275 K. Spectra were processed and analyzed with NMRPipe [171], Sparky, and CCPNMR/Analysis software. External referencing to 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) (for ^{13}C) was done indirectly via the ^{13}C signals of adamantine [97]. For all included spectra, additional experimental details can be found in **Appendix B** (**Table A.3**).

Intra-residue ^{13}C assignments were based on 2D ^{13}C-^{13}C experiments obtained with ^{1}H-^{13}C cross polarization (CP), dipolar assisted rotational resonance (DARR) [35] mixing from 8-100 ms, 83 kHz two pulse phase modulation (TPPM) [98]^{1}H decoupling during acquisition and evolution, and MAS spinning rates between 8-10 kHz. 1D and 2D ^{15}N-^{13}C double CP-based experiments were used to make ^{15}N and connected ^{13}C assignments, complemented with z-filtered TEDOR spectra [172] that circumvent weak ^{1}H-^{15}N CP efficiency for Pro. Side chain ^{13}C assignments for the samples PG1 and pG2 were confirmed using 2D single quantum - double quantum (SQ-DQ) experiments with 1.8 ms and 1.2 ms $SPC5_3$ mixing [173] to generate DQ coherence, respectively.

3.2.2.2 Structural MAS ssNMR Measurements

Carbon proximities were probed via ^{13}C-^{13}C recoupling experiments using 100 ms DARR mixing, or 250-500 ms proton-driven spin diffusion (PDSD) mixing, at 10 kHz MAS. Backbone torsion angle experiments were performed via ^{13}C SQ-DQ NCCN experiments, as previously described [41, 158]. In the NCCN experiments, SPC5 recoupling was used to generate the Cα-CO DQ coherence, using 10 kHz MAS spinning rates, 50 kHz ^{13}C rf power and 100 kHz CW decoupling [174]. The DQ signal was dephased under the ^{15}N-^{13}C dipolar couplings, using variable-time rotational echo double resonance (REDOR) mixing with 50 kHz ^{15}N and ^{13}C rf powers for the respective π pulses.

Side-chain torsion angle experiments were performed using DQ-based HCCH style experiments, in analogy to previous work [43, 175, 176]. SPC5 was used to establish a DQ coherence between the side chain carbons. Variable-time Lee-Goldburg (LG) irradiation on ^{1}H allowed for measurements of the χ_1 and χ_2 dependent dephasing curves. To extract backbone and side chain torsion angles from the NCCN and HCCH experimental results, numerical simulations of the respective pulse sequences were performed with the SPINEVOLUTION program [40]. These simulations were done on multi-core workstations or a departmental 48-core high-performance Linux computer.

3.2.2.3 Quantitative Analysis with SPINEVOLUTION

Quantitative analysis of structural experiments was conducted using SPINEVOLUTION [40]. Each simulation requires the input of a coordinate system and a series of pulses defined by a duration and power level. The signal intensity of the spins is output as a function of time and varied parameters.

The pulse sequence of the NCCN experiment is shown in **Figure 1.7a**. The dephasing curves (**Figure 1.7b**) can be obtained by simulating only pulses on the ^{13}C and ^{15}N channels, including the DQ excitation and reconversion periods on ^{13}C and the REDOR ^{15}N recoupling period on the ^{15}N channel. The coordinate system used for the NCCN simulation was N_i-$C\alpha_i$-CO_i-N_{i+1}, where the N_i-$C\alpha_i$ distance was 1.33 Å, the $C\alpha_i$-CO_i distance was 1.52 Å, the CO_i-N_{i+1} distance was 1.46 Å, the angle defined by N_i-$C\alpha_i$-CO_i was 114.1°, and the angle defined by $C\alpha_i$-CO_i-N_{i+1} was 117.8°. These four nuclei define the torsion angle varied within the SPINEVOLUTION simulation. For a given torsion angle, the signal intensity of the ^{13}C spins was simulated with an increasing REDOR period. Dephasing curves were then simulated for an array of NCCN torsion angles between -180° – 180°. To fit the data, reduced χ^2 was calculated for each of these curves relative to the experimental data, and a reduced χ^2 test was used to find torsion angles that fit within a 90% confidence interval. The experimental error for each data point was estimated by taking the noise level in the 2D spectrum relative to the Gaussian integration of the peak and was considered in the fitting algorithm.

The pulse sequence of the HCCH experiment is shown in **Figure 1.8a**. The portion of the experiment following the ^{1}H-^{13}C CP is simulated including all irradiation periods on both the ^{1}H and ^{13}C channels. The coordinate systems to which the simulations were applied were either Hα-Cα-Cβ-$H\beta_2$ (for χ_1) or $H\beta_2$-Cβ-Cγ-$H\gamma_2$ (for χ_2), where the bonded H-C distances were 1.13 Å, the C-C distance was 1.33 Å, and the angles defined H-C-C were 109.5°. For a given torsion angle, the ^{13}C signal intensity was recorded as a function of increasing LG period up to the length of one rotor period. Interference curves were simulated for an array of HCCH torsion angles between -180° – 180°. The data were fit using a reduced χ^2 test, as described above for the

NCCN experiment. The experimental error in the 1D peak heights was estimated based on the noise level of a spectral region lacking peaks in the respective spectra.

3.2.2.4 ^{15}N R_1 Relaxation Measurements

^{15}N longitudinal R_1 relaxation was measured for backbone and Gln side chain ^{15}N sites using a series of ^{1}H-^{15}N CP experiments incorporating a ^{15}N R_1 relaxation period and 83 kHz TPPM ^{1}H decoupling during acquisition, analogous to earlier work [177]. These measurements were performed at 22 kHz MAS and 600 MHz ^{1}H frequency, conditions where systematic MAS-dependent measurements had shown ^{15}N-^{15}N spin diffusion to be effectively suppressed.

3.2.2.5 Ultrafast MAS ^{13}C and ^{15}N Relaxation Measurements

I acknowledge Jonathan M. Lamley and Dr. Józef R. Lewandowski for measuring ^{15}N $R_{1\rho}$, ^{13}C $R_{1\rho}$ and ^{13}C R_1 relaxation rates at an MAS frequency of 60 kHz using a Bruker 1.3 mm triple resonance probe. These experiments were performed with a Bruker Avance II+ spectrometer operating at a ^{1}H Larmor frequency of 600 MHz. Sample cooling was applied using a Bruker BCU-X variable temperature unit such that the internal sample temperature was held at 296 ± 1 K, measured from the chemical shift of water protons with respect to an internal DSS reference. Relaxation rates were measured from peak intensities in series of 1D spectra, with maximum spin-lock pulse lengths of 300 ms (^{15}N $R_{1\rho}$), 160 ms ($^{13}C'$ $R_{1\rho}$) and 80 ms ($^{13}C_{aliphatic}$ $R_{1\rho}$) and maximum relaxation delay times of 15 s ($^{13}C'$ R_1) and 5 s ($^{13}C_{aliphatic}$ R_1). For each scan, initial ^{1}H magnetization was prepared with a 100 kHz 90° pulse, and subsequently transferred to $^{13}C/^{15}N$ via double quantum cross-polarization of 2.0 ms (^{15}N), 1.4 ms ($^{13}C'$) or 1.0 ms ($^{13}C_{aliphatic}$) with nutation frequencies of ~50 kHz (^{1}H) and 10 kHz ($^{13}C/^{15}N$). Spin-lock nutation frequencies for $R_{1\rho}$ experiments were set to 17 kHz, while for R_1 experiments the 90° pulses

either side of the incremented relaxation delay were set to 100 kHz (^{13}C) or 83 kHz (^{15}N). During acquisition, low-power (15kHz) slpTPPM [178] ^{1}H decoupling was applied.

3.3 RESULTS

3.3.1 The Doubled PolyQ Signature is Present Below and Above the Threshold

Our previous work revealed a seemingly characteristic doubled ssNMR signature (even for a single labeled Gln) in polyQ peptides aggregating via a monomeric nucleus. However, it was also reported that individual labeled Gln residues in fibrils from a shorter $D_2Q_{15}K_2$ peptide lack the doubled NMR signals [22]. Based on NMR study of a peptide with multiple labeled Gln residues, it was found that each residue had only one set of NMR signals, which depended on its proximity to either the N- or C-terminus of the peptide. In an attempt to reproduce and study this potentially polyQ-threshold-dependent phenomenon, we obtained the same peptide featuring a single ^{13}C,^{15}N-labeld Gln, and prepared amyloid-like fibrils for study by MAS ssNMR. **Figure 3.1** shows sections from a ^{13}C-^{13}C 2D spectra comparing the NMR signals of a single ^{13}C,^{15}N-labeled Gln in these $D_2Q_{15}K_2$ fibrils (**Figure 3.1b**) to one in $K_2Q_{30}K_2$ fibrils (**Figure 3.1a**). Despite the fact that in each case only a single Gln is labeled, both spectra unequivocally show two sets of correlated NMR signals. These observations match our previous studies of polyQ amyloids, but contradict the results reported by Schneider et al [22]. Thus, seemingly independent of the length of the polyQ stretch, each single Gln of the amyloid core assumes both of the major conformers (‘a’ and ‘b’) in equal proportions. A minor third peak is seen in $D_2Q_{15}K_2$, which reflects a more mobile conformer ‘c’ that is due to a proportion of sample in which the labeled residue is outside the amyloid core [22, 45].

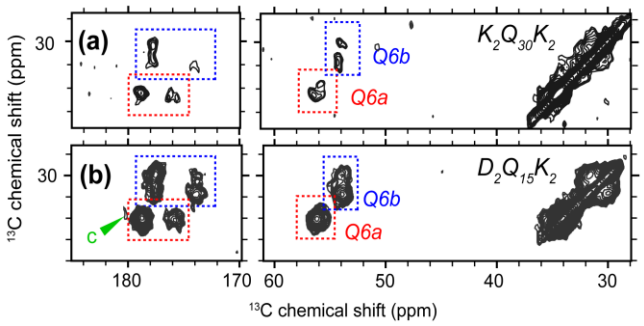

Figure 3.1 Amyloid core signature motif.

Aliphatic to carbonyl (left) and intra-aliphatic (right) regions of 2D ^{13}C-^{13}C DARR spectra of (a) [U-^{13}C,^{15}N-Q6]-$K_2Q_{30}K_2$ fibrils (25 ms DARR mixing) [21] and (b) Q15 fibrils (8 ms DARR mixing) with 10 kHz MAS. The two sets of peaks are highlighted in red (conformer 'a') and blue (conformer 'b'). The single isotopically enriched Gln in both peptide fibrils has both conformers. A third minor conformer, 'c', is also observed for some of the labeled Glns (green).

3.3.2 Kinetics-Modifying Interruptions Do Not Change the PolyQ Core Motif

We have previously reported that the very same signature motif is also seen in several other polyQ-based peptide aggregates featuring a Pro-Gly insertion [60]. This Pro-Gly motif was designed to favor β-hairpin formation in the polyQ segment, which did neither change the aggregation kinetics nor the fibril morphology. Here, we set out to determine the impact of a D-Pro-Gly motif, which is a stronger β-hairpin enhancer and was shown to *enhance* aggregation kinetics. Fibrils were prepared from $K_2Q_{11}pGQ_{11}K_2$ (peptide *pG1*) featuring U-^{13}C,^{15}N-labeling of two consecutive Glns (Q8 and Q9; **Figure 3.2a**). The resulting fibrillar aggregates have a similar morphology to those obtained from undisrupted polyQ (**Figure 3.3**).

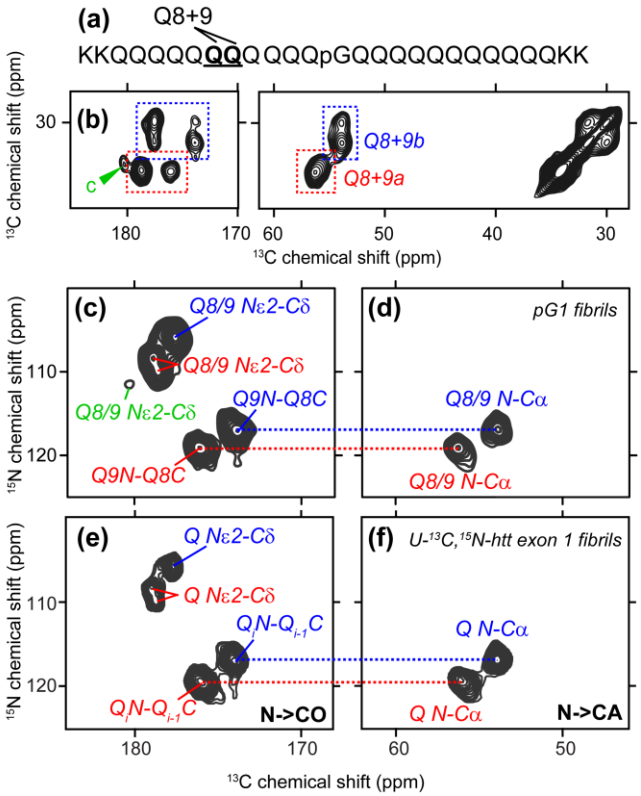

Figure 3.2 Amyloid core conformer connectivity.

(a) Amino acid sequence and labeled sites, Q8 and Q9, of peptide pG1. (b) 2D ^{13}C-^{13}C spectrum with 8 ms DARR mixing of pG1, highlighting signals of Gln conformers 'a' (red) and 'b' (blue) as in ***Figure 3.1****. Aliphatic- carbonyl (left) and intra-aliphatic (right) regions are shown. (c,d)2D ^{15}N-^{13}C spectra of pG1 showing (c) N-CO and (d) N-Cα connectivity. Conformer 'a' signals are labeled in red, conformer 'b' signals in blue, and conformer 'c' signals in green. The Q9a amide bears a cross peak only to Q8a CO, and Q9b amide only show a cross peak with Q8b CO. The same connectivity is seen in U-^{13}C,^{15}N-htt exon 1 fibrils (e,f). The NCO spectrum of the uniformly labeled htt exon 1 fibrils is shown in (e), while its NCA spectrum is shown in (f).*

In 2D ^{13}C-^{13}C spectra (**Figure 3.2b**) these fibrils also feature two sets of NMR signals. Although these correspond to the polyQ signature signals, it is not immediately certain that both Q8 and Q9 each have a doubled signal. However, on the basis of heteronuclear ^{15}N-^{13}C spectra (**Figure 3.2c**) it is clear that this indeed once again is the case. Moreover, the connectivity between these

conformers is apparent: in any particular peptide, the two neighboring residues Q8 and Q9 are either both in conformation ‘a’ or both in conformation ‘b’.

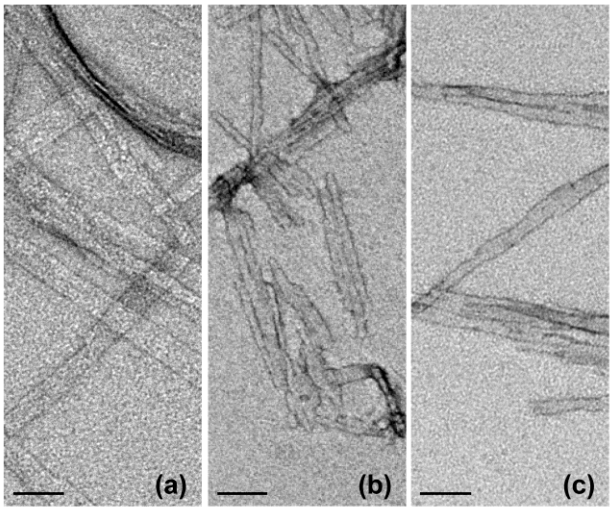

Figure 3.3 Negatively stained EM of polyQ aggregates.

(a) uninterrupted polyQ, $D_2Q_{15}K_2$. (b,c) D-Pro and L-Pro containing polyQ peptides [25], (b) $K_2Q_{10}pGQ_{11}K_2$ and (c) $K_2Q_{10}PGQ_{11}K_2$, respectively. The scale bars indicate 50 nm.

Application of similar experiments to fibrils formed uniformly-^{13}C, ^{15}N- labeled htt exon 1 reveals identical inter-residue connectivities (**Figure 3.2e-f**). Earlier work suggest that these observations also apply to other polyQ fibrils [22]. Thus, it appears that these ‘a’-to-‘a’, ‘b’-to-‘b’ connectivities between neighboring Gln residues are conserved features of different polyQ-based amyloid cores, ranging from simple polyQ model peptides to disease-relevant expanded htt exon 1 fibrils (**Figure 3.2**).

3.3.3 The PolyQ Amyloid Core is Based on β-sheets

Although it is widely assumed that polyQ amyloids are β-sheet-based, like other amyloids, there are various reports that challenge this assumption. We previously reported that the polyQ fibril chemical shifts are indicative of β-sheet structure [21]. On the basis of the unequivocal assignment that connected Gln residues tend to feature the same chemical shifts, we can also use the software TALOS+ (Torsion Angle Likelihood Obtained from Shift and sequence similarity) [39] to constrain the backbone torsion angles of each of the two polyQ amyloid core conformers (**Figure 3.4b, d**). This analysis indicates that both conformers feature backbone torsion angles characteristic of a β-sheet secondary structure. Nonetheless, previous studies that have argued in favor of α-sheet-based polyQ aggregate structure have suggested that chemical shift analysis may not be an unequivocally reliable approach [168, 179]. To directly address this concern, we use MAS ssNMR backbone torsion angle measurements that do not rely on chemical shifts, but rather on the relative orientations of dipolar coupling vectors. The Q8/Q9-labeld pG1 fibrils were submitted to NCCN torsion angle measurements (see methods). This experiment is sensitive to the relative orientation of the N-Cα and CO-N dipoles and thus reports on the ψ torsion angle for each of the Gln conformers [41]. **Figure 3.4a and c** show the NCCN results as a plot of the normalized ^{13}C DQ signal dephasing due to a 0.8 ms REDOR period. Experimental results are shown in red and blue for conformers 'a' and 'b', respectively, with the range of error shown in pale colors. Black curves show the theoretical S/S_0 values for all ψ values (**Figure 3.4a and c**). These results constrain conformer 'b' to 152 +/- 2°. This value is indicative of β-sheet structure and is explicitly inconsistent with α-sheet structure, which has 92 ± 28° or -49 ± 4° [168]. The NCCN curve for conformer 'a' is somewhat ambiguous, as it can be

fit to multiple values. However, we note that only the ψ of 134 +/- 2° is consistent with the TALOS+ data (**Figure 3.4b**). We also note that an α-sheet structure would imply that alternating residues feature different backbone conformations, which would most likely result in patterns of alternating chemical shifts (unlike the data in **Figure 3.2**). Thus, our results support the idea that the two ssNMR signals reflect two types of β-strands, with all the Gln within each strand type featuring the same chemical shifts.

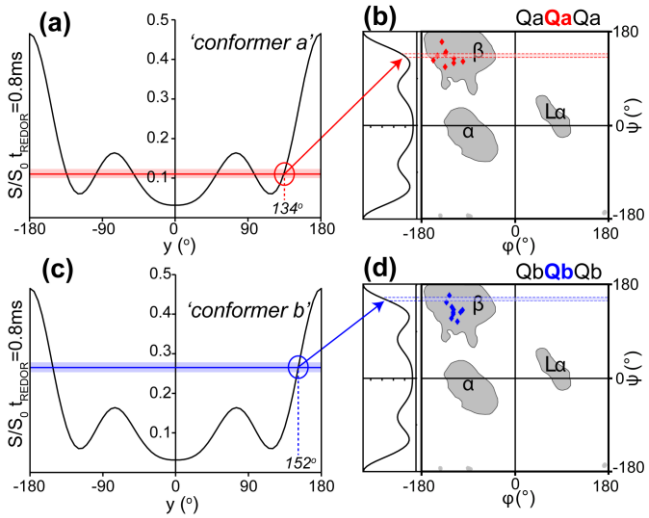

Figure 3.4 Constraining backbone torsion angles.

(a,c) NCCN signal intensity of Q8 Cα and CO at t_{REDOR}=0.8 ms due to ^{15}N recoupling (S/S_0) is plotted for (a) conformer 'a' (red solid line) and (c) conformer 'b' (blue solid line). Pale colors represent error. In each panel, the black curve is the theoretical S/S_0 for all ψ angles. (b,d) Backbone torsion angles for the ten known structures closest in chemical shift and primary sequence by TALOS+ analysis to (b) a tri-peptide of Gln with chemical shifts of conformer 'a' and (d) a tri-peptide of Gln with chemical shifts of conformer 'b' are plotted as red and blue diamonds, respectively, on the Ramachandran plot. Highlighted regions of ψ are those where NCCN and TALOS+ results converge.

3.3.4 The PolyQ Conformers have Differing Dynamic Properties

MAS ssNMR was also used to characterize the dynamics that characterize each of these two β-strand types. ^{15}N longitudinal relaxation (R_1) were applied to the single U-^{13}C,^{15}N-labeled residue Q10 in $K_2Q_{11}PGQ_{11}D_2$ fibrils [25]. These measurements were conducted at 600 MHz ^{1}H frequency and 22 kHz MAS, where ^{15}N-^{15}N spin diffusion is sufficiently suppressed. In these experiments, an R_1 relaxation delay was incorporated into a ^{1}H-^{15}N CP 1D experiment, and the decay ^{15}N signal intensity was observed with increasing relaxation delay time. The R_1 values were calculated from these single exponential decay curves and are given on the bar graphs in **Figure 3.5b**. Our collaborators at the University of Warwick, Dr. Józef Lewandowski and Jonathan Lamley, measured the ^{13}C R_1 and $R_{1\rho}$ and ^{15}N $R_{1\rho}$ relaxation rates at ultrafast MAS at 60 kHz, shown in **Figure 3.5a, c, and d**. For all ^{15}N and the vast majority of ^{13}C sites, conformer 'a' has a significantly higher R_1 and $R_{1\rho}$ than those of conformer 'b' (**Figure 3.5**). These data indicate that β-strands featuring conformer 'a' are less stable than those featuring conformer 'b.' This applies to both the side chains and the backbones, and is consistent with previously published ^{15}N R_1 data [45]. However, we cannot exclude that this is reflective of different ^{1}H environments or proximities of ^{1}Hs for the ^{15}N sites of the distinct conformers.

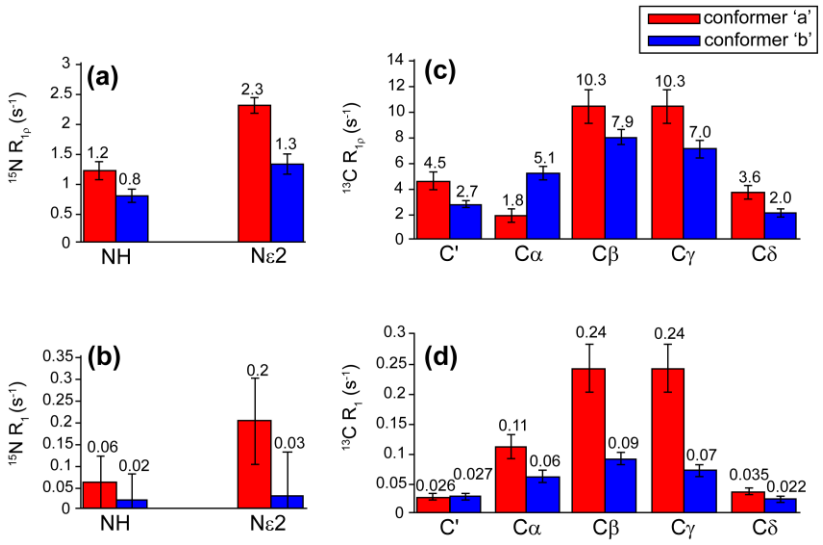

Figure 3.5 ^{13}C and ^{15}N relaxation measurements of [U-^{13}C,^{15}N-Q10]-$K_2Q_{11}PGQ_{11}D_2$ fibrils.

(a) $^{15}N\ R_{1\rho}$, (b) $^{15}N\ R_1$, (c) $^{13}C\ R_{1\rho}$, and (d) $^{13}C\ R_1$. $^{15}N\ R_1$ was measured at 22 kHz MAS, while the others were at 60 kHz MAS. In each graph, conformer 'a' is depicted as red bars and conformer 'b' as blue bars. For most sites within the enriched Gln, conformer 'a' shows a higher relaxation rate.

3.3.5 Both PolyQ Conformers Feature Extended Side Chain Conformations

As the major difference in chemical shift between conformers 'a' and 'b' is characteristic Cβ and Cγ chemical shifts, it may seem that structural differences would arise from substantial differences in the χ_2 side chain angles. To probe this directly, we used HCCH experiments that examine the orientation of the HβCβ – CγHγ dipolar vectors [43]. The observed Cβ-Cγ HCCH interference patterns for conformers 'a' and 'b' are not significantly different. Moreover, these dephasing curves are only consistent with χ_2 angles near 180°, implying an extended side chain formation. This is true for both conformers of Gln in the amyloid core of both simple polyQ (**Figure 3.6d**) and those with pG inserts (**Figure 3.6b**). This experiment also reports on the χ_1

angle based on the HCCH interference pattern of the Gln Cα/Cβ signals. The χ_1 angles for the two conformers are significantly different in both uninterrupted polyQ and with incorporation of pG inserts, showing clear structural differences between conformers 'a' and 'b' (**Figure 3.6a and c**). We find the best fit χ_1 angle for conformer 'a' to be -65^o and that for conformer 'b' to be 55^o independent of the polyQ context. The simulated curves of these χ_1 torsion angles are shown in **Figure 3.6a and c**. However, reduced χ^2 analysis cannot exclude other rotomeric states based on these data alone. Nonetheless, these data together with ab initio calculations (not shown) seem most consistent with χ_1 angles of ~ -60^o and 60^o for the two conformers. Regardless, it is clear that the χ_1 values of the 'a' and 'b' conformers are distinct, while the χ_2 angles are similar (near 180°).

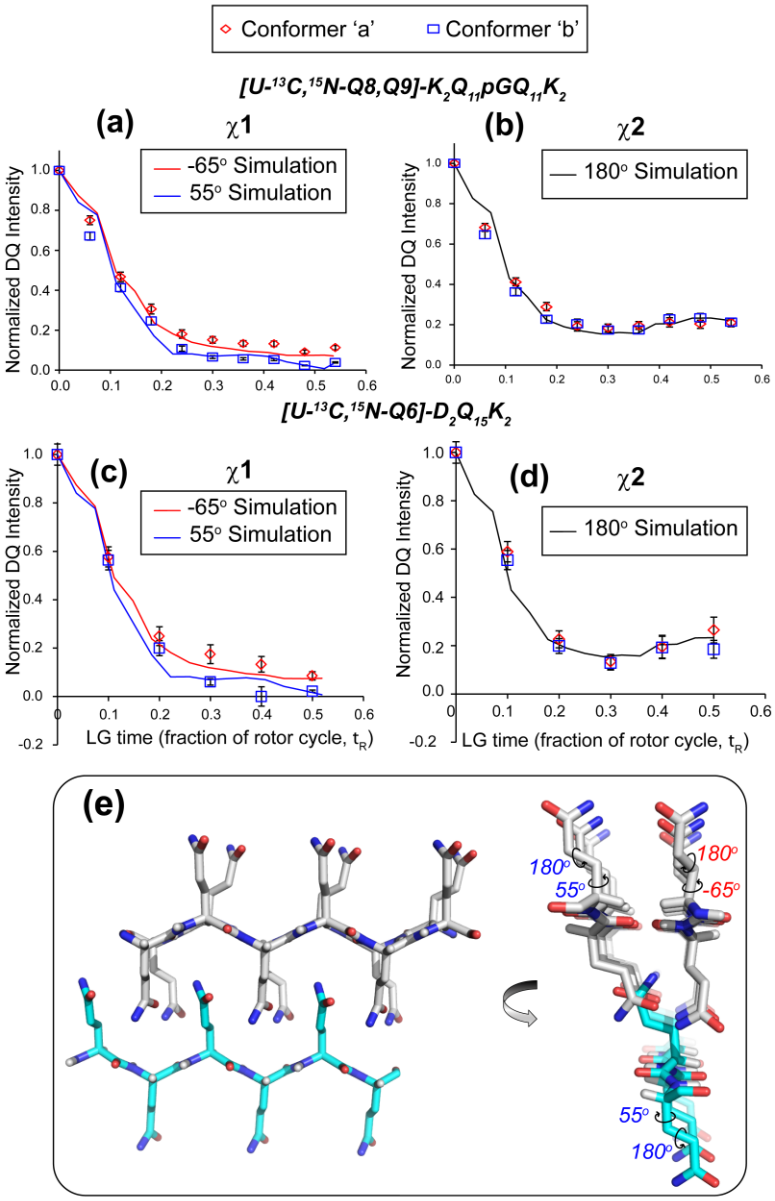

Figure 3.6 Constraining Gln side chain torsion angles.

HCCH data points for Gln conformers 'a' and 'b' are plotted as red squares and blue diamonds, respectively. (a,c) Normalized DQ intensity as a function of Lee-Goldburg (LG) decoupling time for Cα of (a) pG1 and (c) Q15 fibrils. This interference pattern is indicative of the χ_1 angle. Simulated data are plotted as solid red and blue curves, respectively. (b,d) Normalized DQ intensity as a function of LG decoupling time for Cβ and Cγ of (b) pG1 and (d) Q15 fibrils, giving the χ_2 angle. Both conformers best fit to a χ_2 angle of ~ 180°; the simulated curve is plotted as

the solid black line. (e) A model of polyQ with extended side chains (χ_2= 180^o) and full strands of conformer 'a' (χ_1= -65^o) and 'b' (χ_1= 55^o) χ_1 torsion angles.

3.3.6 Lack of Peak Doubling Shows Conformers 'a' and 'b' Within a Single Peptide

The previous data may appear consistent with fibrils that are made up of peptides that in their entirety adopt one of the two conformers, or even having entire β-sheets or fibrils formed from one conformer. In the course of studying polyQ peptides featuring Pro-Gly-based insertions, we have encountered at least one scenario in which the peak doubling is disrupted, allowing an intriguing insight into the structural roles of the two conformers. Based on labeling of residues away from the pG or PG insertions, we showed that the amyloid core is indistinguishable from uninterrupted polyQ [25]. We also prepared fibrils from peptides in which ^{13}C,^{15}N-labeled Gln were purposely incorporated close to the sites of mutation. Assignment spectra of pG2, featuring isotopically labeled residues [U-^{13}C,^{15}N-Q13,G15,Q16], are shown in **Figure 3.7**. Heteronuclear experiments were used to distinguish Q16 from Q13 (only the former is preceded by a labeled residue), including the application of a CONCX experiment that results in selective polarization of the Q16 Cα (**Figure 3.7**).

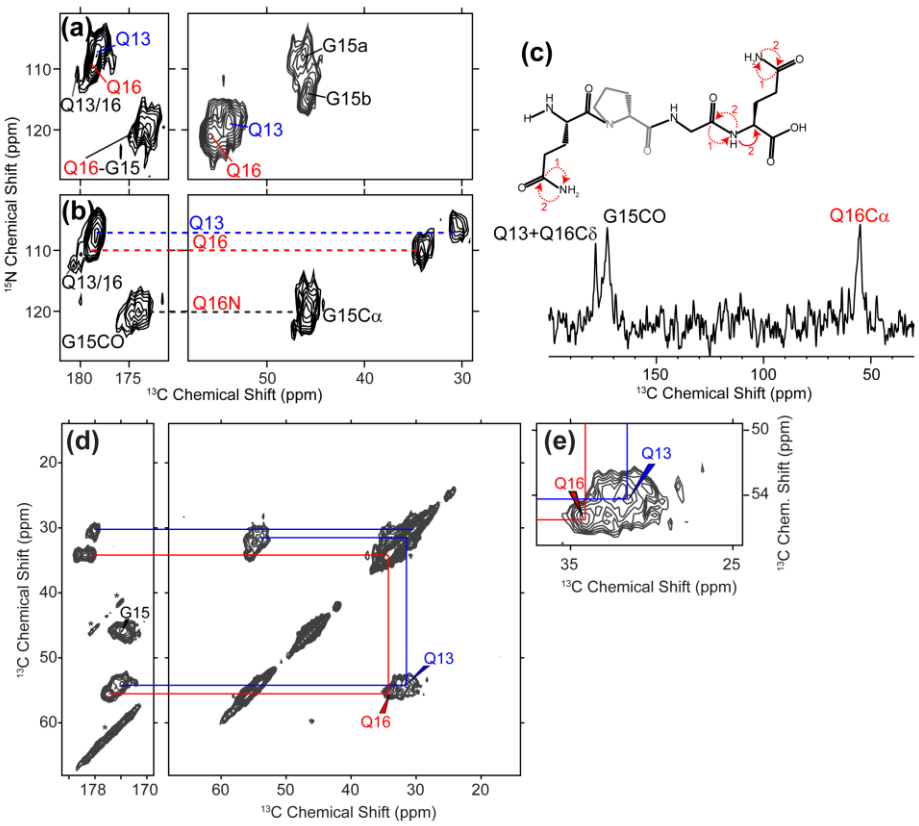

Figure 3.7 Chemical shift assignment experiments for Q13, G15, and Q16 in pG2.

(a,b) ^{15}N-^{13}C 2D spectra of pG2. a) N-CO correlations are shown on the left and N-Cα on the right. Q13 and Q16 resonances are highlighted in blue and red, respectively. b) NCOCX. N-CO region is shown on the left and N-$C_{aliphatic}$ is shown on the right. Dashed lines highlight the correlations to the indicated ^{15}N. c) CONCX ^{13}C 1D spectrum, which gives the Q16 Cα assignment at 55.4 ppm (red) used to differentiate Q16 from Q13. Since the D-Pro N is not isotopically labeled (gray), Q13 Cα is not accessible through the CO-N-Cα magnetization transfer pathway, illustrated by the red arrows on the chemical structure of QpGQ. (d,e) ^{13}C-^{13}C 2D MAS ssNMR spectra on fibrils from pG2 (8ms DARR). d) Aliphatic-carbonyl (left) and intra-aliphatic (right) spectral regions are shown. e) Zoomed-in Cα-Cβ cross peaks to highlight key signature cross peaks of isotopically labeled Gln.

Key aliphatic cross peaks of Q13 and Q16 are shown in **Figure 3.8a**. Strikingly, Q16 does not feature both sets of chemical shifts, but has chemical shifts that match conformer 'a'. In contrast, Q13 strongly resembles the other conformer (i.e., 'b') of the doubled polyQ signature.

The bar graphs in **Figure 3.8d** show the similarity in chemical shifts between pG turn-flanking Gln and the signature motif conformers in an amyloid core Glns (Q10). The chemical shifts for isotopically labeled residues in all polyQ-based fibrils in this study are given in **Table A.4** in **Appendix B**. Thus, in the pG constructs, Gln residues directly surrounding the pG motif manage to participate in the β-sheet amyloid core, but each takes on only one conformer. This demonstrates that these two conformers must therefore be part of a single peptide, and are not due to distinct fibrils.

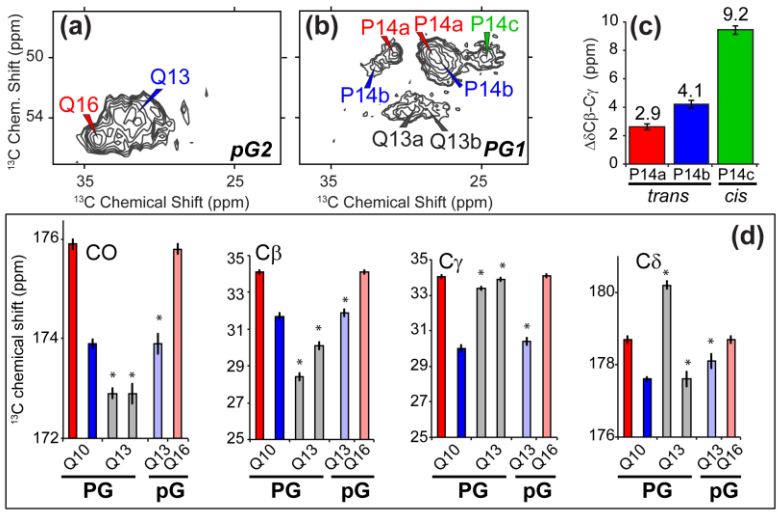

Figure 3.8 Chemical shift assignments in pG and PG-interrupted polyQ.

(a,b) Intra-aliphatic regions from ^{13}C-^{13}C 2D MAS ssNMR spectra of (a) pG2 (8 ms DARR mixing) and (b) PG1 (10 ms DARR mixing) highlighting the different Gln Cβ-Cγ chemical shifts. These data were obtained at 600 MHz ^{1}H frequency and 10 kHz MAS. The three P14 conformations in PG1 are highlighted in color in (b). (c) Bar graphs show the difference between Cβ and Cγ chemical shifts in the three P14 conformers, color coded as in (b). P14c has a much larger difference between these shifts, indicative of a cis-conformation. (d) Bar graphs show the CO, Cβ, Cγ, and Cδ signature chemical shifts of conformers 'a' and 'b' in Q10, in the amyloid core, of [U-^{13}C,^{15}N-Q10]-$K_2Q_{11}PGQ_{11}D_2$ and those in Q13 and Q16 in pG2. Q13 and Q16 chemical shifts bear strong resemblance to the conformer 'b' and 'a' chemicals shifts of the signature motif, respectively. Chemical shifts of the major conformers

of Q13 in PG1 are also show (gray), which are unlike the signature 'a' and 'b' conformers. Asterisks indicate Gln residues adjacent to the Pro.

The above turn motif contrasts notably with the L-Pro-Gly-based one, as seen in $K_2Q_{11}PGQ_{11}D_2$. These peptides aggregate more slowly than pG, but follow kinetics similar to undisrupted polyQ peptides of similar length, as reported previously [25]. We examined peptides with isotopic ^{13}C,^{15}N-labeling in Q13, P14, and G15 (PG1) (**Table 3.1**), and obtained chemical shift assignments of the labeled residues. Here, Q13, adjacent to the Pro, lacks these signature shifts of the polyQ core. It features at least two different conformations with their own distinct sets of chemical shifts, which do not match either conformer of the polyQ core signature motif (**Figure 3.8b and d**), but are still characteristic of β-sheet structure. This is suggestive of a position outside the amyloid core, similar to the final Gln preceding the prolines in htt fragment fibrils [45]. This is supported by ^{15}N longitudinal relaxation measurements. The backbone nitrogens have R_1 values from ~ 0.02- 0.2 s^{-1}, suggesting a relatively rigid backbone structure in the turn. However, the Gln Nε2 sites seem to show the fastest relaxation (**Figure 3.9**), indicating increased solvent accessibility relative to the amyloid core. Thus, the L-Pro-Gly motif has a qualitatively different impact on its polyQ environment.

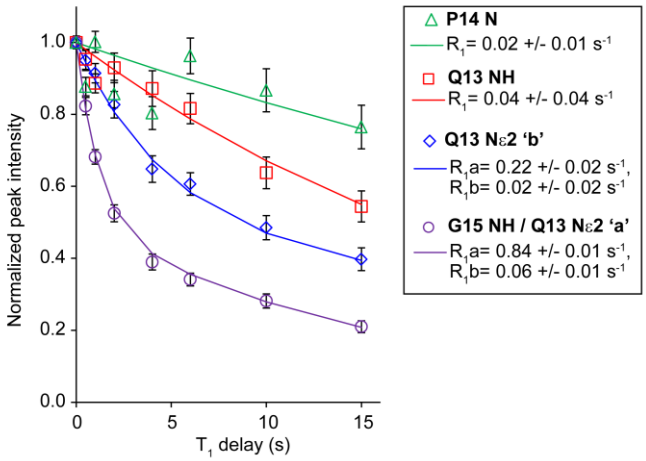

Figure 3.9 ^{15}N longitudinal relaxation for PG1 fibrils.

1D ^{15}N spectra were acquired at 22 kHz MAS and 275 K. R_1 fits are shown as solid lines for each ^{15}N resonance. Overlapping signals of G15 and Q13 Nε2 are indicated, as well as double exponential fit parameters where appropriate.

3.3.7 Insights from the Pro-Gly Insertions

By inspection of the labeled L-Pro and Gly residue signals, we see that the Gly and Pro residues are heterogeneous: each features three distinct sets of chemical shifts. The three Pro conformers, 'a', 'b', and 'c', constitute approximately 20, 40, and 40% of the sample, based on relative peak intensities. The first two (i.e. 60% of the sample) have Cβ and Cγ shifts that indicate a trans Gln-Pro peptide bond. The third conformer (40% of the total) has a cis configuration, as indicated by a significantly larger Cβ-Cγ chemical shift difference (**Figure 3.8c**) [180]. ^{13}C-^{13}C correlation experiments also show G15Cα-Q13 cross peaks, e.g. between G15Cα and Q13 Cα at 500 ms PDSD mixing (**Figure 3.10a**), for both of the Gln conformers.

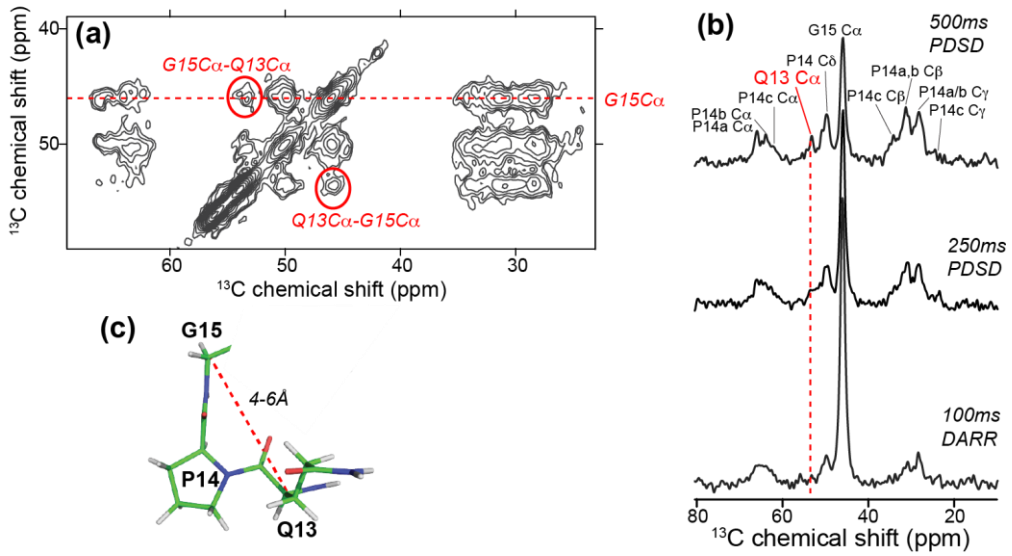

Figure 3.10 ^{13}C-^{13}C transfer experiments.

*(a) Intra-aliphatic region from a ^{13}C-^{13}C 2D with 500 ms PDSD mixing. The red dashed line highlights the cross peaks with G15 Cα, including Q13a and b Cα, and represents the extracted 1D slices shown in (b). (b) ^{13}C 1D extracts of cross-peaks with the G15 Cα from (top-bottom) 500 ms PDSD, 250 ms PDSD, and 100 ms DARR spectra (acquired at 800 MHz ^{1}H frequency). Red dashed lines highlight Q13 resonances. Individual conformations are not indicated where they cannot be separated in the 1D. Corresponding 2D data were acquired with 8 kHz MAS at 275 K for 28 hrs. (500 ms PDSD), 26.5 hrs. (250 ms PDSD), and 24 hrs. (100 ms DARR) (**Figure 3.11**). (c) Illustration of a β-II turn highlighting the observed i → i+2 contacts.*

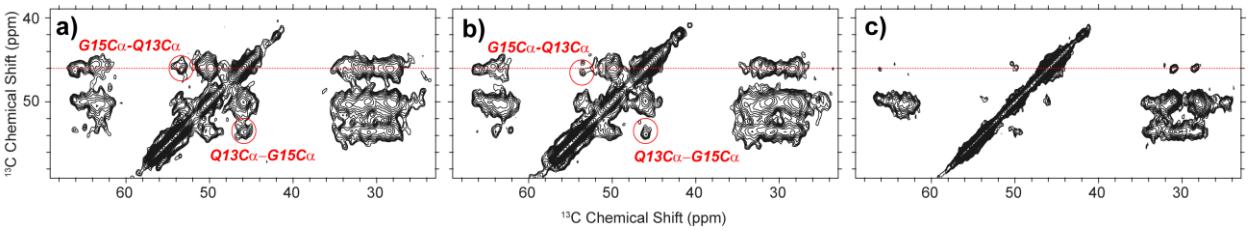

Figure 3.11 Series of long distance ^{13}C-^{13}C transfer experiments via PDSD.

2D ^{13}C-^{13}C spectra of pG2 with a) 500 ms PDSD, b) 250 ms PDSD, and c) 100 ms DARR. The red dashed line designates the G15 Cα transfers, including to the Q13a and b Cα (red circles) with 250 ms PDSD (b). These peaks are not observed with 100 ms DARR mixing (c). Even with 250 ms PDSD, a contact is observed between G15 Cα

and Q13 Cα, which is two residues away (indicated by the red circles). This indicates that the Cαs are less than ~6 Å apart in space.

Thus, G15 and Q13 have to be close in space (<6Å Cα-Cα distance). This is consistent with a β-type turn, but not with an extended β-strand, where the Cα carbons of residues i and i+2 are ~7 Å apart. Chemical shift comparisons, TALOS+ analysis, and NCCN experiments on the L-Pro-Gly samples (**Figure 3.12, Figure 3.13**, and **Table 3.2**) are most consistent with backbone torsion angles that match a β-II turn conformation for both trans-Pro conformations and a type VI-b β-turn for the cis-Pro conformation. These turn structures tend to involve at least four residues, rationalizing the inability of the flanking Gln to partake in the polyQ core.

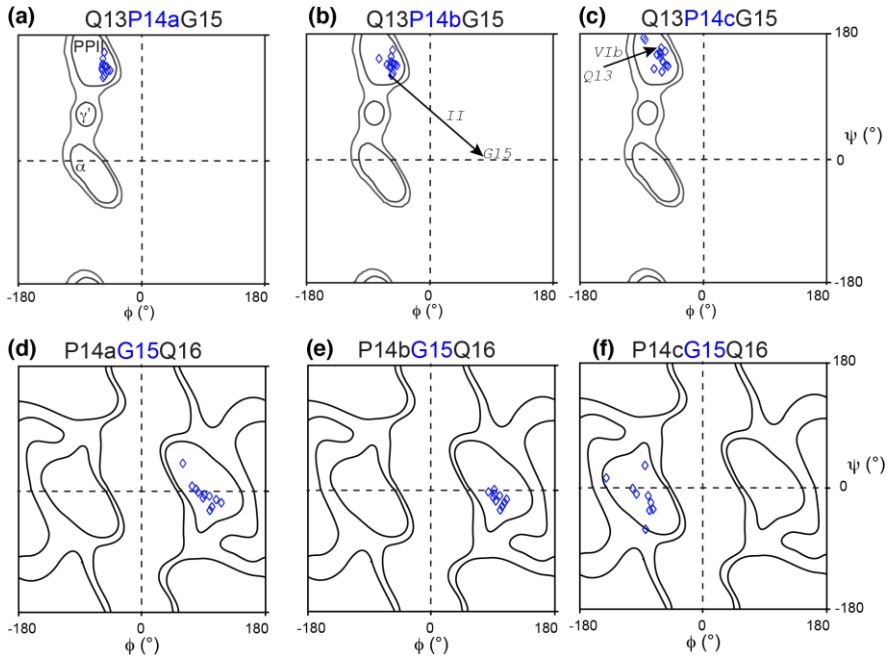

Figure 3.12 Backbone torsion angle analysis of the Pro-Gly motifs.

(a-c) TALOS+ [39] data plotted as open diamonds on the Pro Ramachandran plot [181] for the three distinct sets of resonances. Note that the 'c' form features the cis-Pro. (d-f) Analogous results for Gly15 overlayed on the Gly Ramachandran plot [181] when combined with the presumed Q16 chemical shifts. The employed algorithm

considers a three-residue segment (shown), but only yields constraints on the central residue. Arrows and italicized text in (b,c) indicate the canonical angles for residues 2 and 3 in type II and VI-β turns [181].

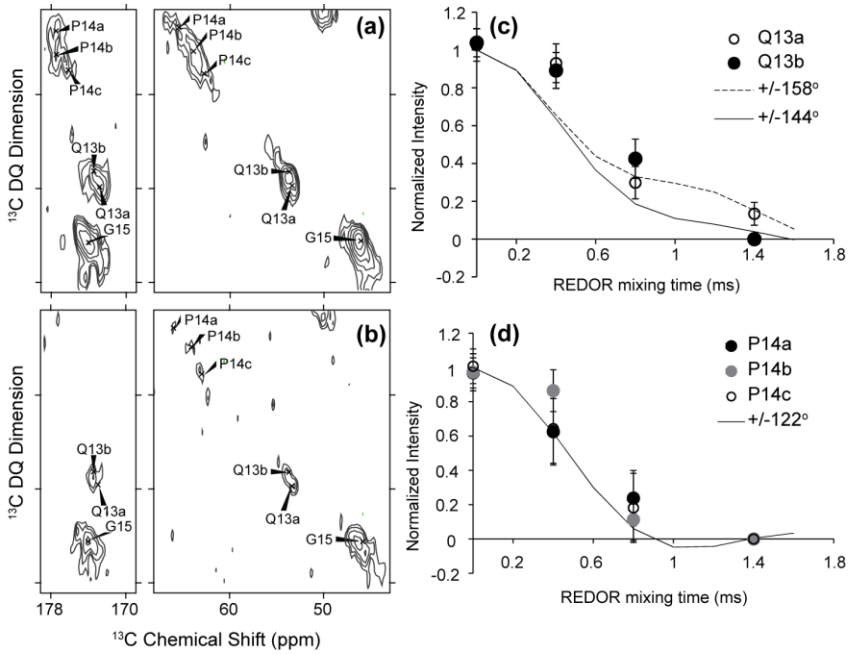

Figure 3.13 NCCN data for Q13 and P14 conformers.

(a,b) Assignment of 2D SQ-DQ NCCN spectrum of PG1. These data were acquired with a 0.8 ms REDOR period with 0 kHz (a) or 50 kHz (b) ^{15}N π pulses. The carbonyl (left) and aliphatic (right) regions are shown. (c) NCCN dephasing curve of Q13 conformers (Q13a- open circles, Q13b- filled circles). (d) NCCN dephasing curve for P14 conformers (P14a- filled black circles, P14b- filled gray circles, P14c- open circles). Data points for each REDOR mixing time represent the average volume of the corresponding CO and Cα peak with ^{15}N pulses applied normalized to ^{15}N pulses off (e.g. peak volumes of (b) normalized to peak volumes of (a)). Each data point required two NCCN experiments with experimental times of ~24 hrs each. In each case, lines are simulated curves for the indicated ψ torsion angle fits.

Table 3.2 Backbone ψ torsion angles constrained by TALOS+ and compared to NCCN data fits.

TALOS+ constraints are not made on Q13 due to lack of shift assignments for the preceding residue. The 95% confidence interval is given in parentheses. Some residues fit to multiple angles, in which cases the best fit is listed first.

Residue	**TALOS+ ψ (°)**	**NCCN |ψ| (°)**
Q13a	*	158 (0-180)
Q13b	*	72, 144 (<160)
P14a	136 +/- 10	70, 122 (<160)
P14b	137 +/- 9	70, 122 (<160)
P14c	148 +/- 14	70, 122 (<160)
G15a	-7 +/- 15	*
G15b	-12 +/- 9	*
G15c	-13 +/- 26	

3.4 DISCUSSION

3.4.1 Peak Doubling is Universal in the PolyQ Amyloid Core

We have shown that fibrils formed by short polyQ peptides, expanded polyQ in context of htt exon 1, and Pro-Gly-interrupted polyQ peptides all faithfully reproduce the same polyQ ssNMR signature. This follows our earlier work showing the same for Q_{30} peptides and synthetic N-terminal htt fragments [21]. Work by the Baldus group [22] has shown similar results, but disagreed with one key feature of our polyQ signature, by arguing that the peak doubling is not universal. This was based primarily on their analysis of Q_{15} fibrils featuring multiple labeled Gln residues. We have shown here that Q_{15} fibrils featuring just a single labeled Gln clearly have the doubled signals also seen in longer polyQ (**Figure 3.1**). It is in principle possible that this reflects the formation of a different fibril polymorph, as is common for other amyloids. However, thus far there is little evidence of analogous polymorphism in polyQ. Rather, based on an inspection of the published Q_{15} spectra, we suggest that those data do feature cross peaks consistent with the peak doubling, but that this feature was not identified due to the complexity and peak overlap associated with the multi-residue labeling scheme. Thus, our data suggests that the peak doubling is a universal feature of Gln residues within the polyQ amyloid core. It is seen for polyQ peptides aggregating via a monomeric as well as multimeric nucleus, as well as for various fragments of htt (including exon 1). Between these different constructs, there is a wide range of aggregation kinetics and mechanisms, which in the htt fragments includes α-helix rich oligomers [21].

Through several MAS NMR experiments, we have also provided further insight into the underlying structural features that rationalize the seemingly universal polyQ-amyloid signature. Using site-specifically labeled samples, we showed that neighboring residues feature the same NMR signals. This is in line with similar observations on extensively labeled polyQ peptides [22]. We also found that the same is true for fibrils formed by htt exon 1 (**Figure 3.2**), which undergoes aggregation in HD.

3.4.2 PolyQ Amyloid Features β-sheets, Not α-sheets

One implication of the observed inter-residue connectivity is that it contradicts one key feature of α-sheet structures, which have been suggested for polyQ [168, 182-184]. Unlike α-helix and β-sheet structure, α-sheets are not characterized by a single repeating set of backbone torsion angles. Instead, alternating residues are predicted to occupy different parts of the Ramachandran plot (α_L and α_R regions). This would imply two distinct sets of chemical shifts, as observed for polyQ, but these would occur in an alternating pattern. Further evidence for β-sheet structure, and against α-sheet conformations, was obtained from chemical shift analysis and backbone torsion angle measurements. Combined, it is clear that the polyQ amyloid core is assembled from β-sheets that contain two distinct types of β-strands. Each of these β-strands is characterized by a distinct chemical shift pattern. This is reminiscent of the analysis by Schneider et al., except that they attributed the two major conformations to Gln s occurring in distinct segments of the peptide: they interpreted their data to show that the N- or C-terminal parts of a $D_2Q_{15}K_2$ adopt conformer ‘a’ and ‘b’, respectively. Our data, here and previously,

clearly show that this cannot be the case, as each individual labeled residue consistently is found in *both* β-strand types.

3.4.3 Amyloid Core Gln Have Extended Side Chains Consistent with Interdigitation

Thus, polyQ amyloid is assembled from two distinct types of β-strands that manifest themselves as well-separated peaks in our NMR spectra. We performed MAS NMR measurements that probe the structural and motional features of each of these β-strands. We have used HCCH-style experiments to probe the χ_1 and χ_2 torsion angles of the Gln residues, finding that the Gln side chains appear to adopt extended conformations ($\chi_2 \sim 180°$). This argues against models of polyQ aggregates featuring bent side chains [20]. Extended side chains are likely indicative of interdigitating zipper-like side chains, which are featured in multiple existing models of polyQ amyloid [19, 26, 185]. However, upon close inspection, these models predict alternating chemical shifts within each β-strand, lack an anti-parallel β-sheet structure, or have χ_1 and χ_2 angles that conflict with our experimental data.

3.4.4 Dynamics and Stability of the PolyQ β-strands

We also employed various NMR experiments at moderate and ultra-fast MAS rates, which showed that there is clearly a difference in the dynamics of the two conformers. The R_1 and $R_{1\rho}$ measurements show that both conformers are relatively rigid, in backbone as well as the side chains. Interestingly, comparing the two β-strand types, conformer ‘a’ consistently has higher relaxation rates throughout the backbone and side chain sites. Although it may be

tempting to attribute this to an increased exposure to solvent, previous measurements of solvent exposure have not shown this to be the case [21, 22, 45]. We hypothesize that the underlying cause for increased dynamics in the a-type β-strand, is that it reflects a difference in the stability or strain of the side chain structures in the two types of β-strands. This may help explain one of the few cases where the two conformers deviated from the typical 1:1 ratio, near the boundary of the polyQ amyloid core in htt N-terminal fragment fibrils. The penultimate residue of the polyQ domain was found to have a lower population with the type-a chemical shifts, with a concomitant increase of signals indicative of Gln outside the amyloid core proper [45]. This seems consistent with a reduced stability of conformer 'a', in line with the observed differences in dynamics [45].

3.4.5 Pro-Gly Insertions and β-hairpin Formation

L-Pro-Gly insertions into polyQ have been employed in many studies, which commonly assume they end up in β-turn-type structures in β-hairpins. However, there has been some debate whether this is actually the case. One account has suggested that polyQ fibrils assemble in a β-arcade structure without β-hairpins, and that introduced Pro residues get incorporated into the β-strands [22]. Although the latter is not unprecedented [172, 186], our current data do not support this kind of arrangement. We previously showed that PG-type peptides maintain a typical amyloid core signal away from the turn [25]. Here we probed the local consequences of the PG insertion and found that the neighboring Gln is unable to participate in the amyloid core. Thus the local impact of the L-Pro is somewhat reminiscent of the impact of the prolines in htt N-terminal fragments [45]. The PG-interruptions form a rigid and compact turn structure with a three-fold structural heterogeneity, reflected for instance in an approximately 3:2 ratio of trans-

to cis-Pro isomers. The dominant trans-L-Pro conformers have chemical shifts [187, 188] and backbone torsion angles that are indicative of a β-II turn structure. Note that type II β-turns have i+1 ϕ,ψ angles of -60°, 120°, respectively[189], commonly feature Gly in a central position, and are particularly common for Pro-Gly motifs [190]. The ~40% population of cis-L-Pro seems to match the features of a type VI-b β-turn [181, 191]. As expected for a type VI turn, Q13 has β-type secondary shifts even though it clearly lacks the characteristic NMR signals of the amyloid core. All these L-Pro-Gly turn structures are multi-residue turns that would predict and necessitate a disruption of the surrounding Gln residues, as observed. The structural heterogeneity of the turn is reminiscent of PG turns in other protein contexts [192] and previous ssNMR work on elastin-like peptides [187, 188].

One striking feature of the pG sample measurements is that here the Gln residues flanking the pG motif *do* have resonance frequencies that are very similar to the polyQ amyloid chemical shifts. Thus, this is indicative of a very tight two-residue turn structure. This fits well with prior structural studies on pG-containing polypeptides [193-195] that show a very tight "minimal" two-residue turn structure with flanking residues partaking in the strands of a β-hairpin. Indeed, previous uses of pG were often inspired by its expected ability to adopt compact 'mirror image' turns not accessible to its L-Pro counterpart [167, 196]. Thus, our data suggest that the pG-motif in a polyQ segment also forms such a two-residue "mirror image" turn, which allows neighboring Gln residues to participate in typical polyQ β-strands without much disturbance of the polyQ fibril architecture. These qualitative differences between the L- and D-Pro-Gly motifs seemingly provide a rationale for the differing effects on aggregation kinetics and in particular fibril stability [25].

3.4.6 Implications for Fibril Structure and Formation

The current data provide a number of new clues regarding the origins of the doubled polyQ NMR signal. To our surprise, the D-Pro-Gly insertion resulted in a thus-far unique local cancelation of the peak doubling, as Q13 only adopts the 'b' conformer, whereas Q16 features only the 'a' conformer. This provides further unequivocal evidence that these two conformers cannot be part of co-mixing polymorphs, but rather together assembly into a "composite" fibril structure [177]. Moreover, we interpret this as providing an intriguing new insight into the relationship between the conformers, which also elegantly rationalizes their typically 1:1 ratio. The experiments probing the structure of the Pro-Gly insertions indicate their involvement in different types of β-turns. Similar to pG motifs in other proteins, they also appear to form two-residue turn structures in polyQ. Thus, the two flanking Gln residues, each featuring one of the two polyQ conformers, would be positioned to hydrogen-bond to each other at the very end of the antiparallel β-strands of the β-hairpin. We hypothesize that this defines the natural deposition of the two conformers. Hydrogen bonding of both the Gln backbone and side chains between the antiparallel β-strands necessitates the two strands to adopt distinct structures. The molecular model in **Figure 3.6e** illustrated a combination of backbone and side chain conformations that is consistent with our NMR measurements that assumes χ_1 angles of -65° and 55°. Although this is a preliminary model, it illustrates how the side chain conformations of the two β-strands are complementary in an antiparallel arrangement, but would experience steric clashes in other arrangements (see also **Figure 3.15** below). This type of antiparallel strand-strand complementarity is also manifest in several of the previous structural models of antiparallel polyQ amyloid.

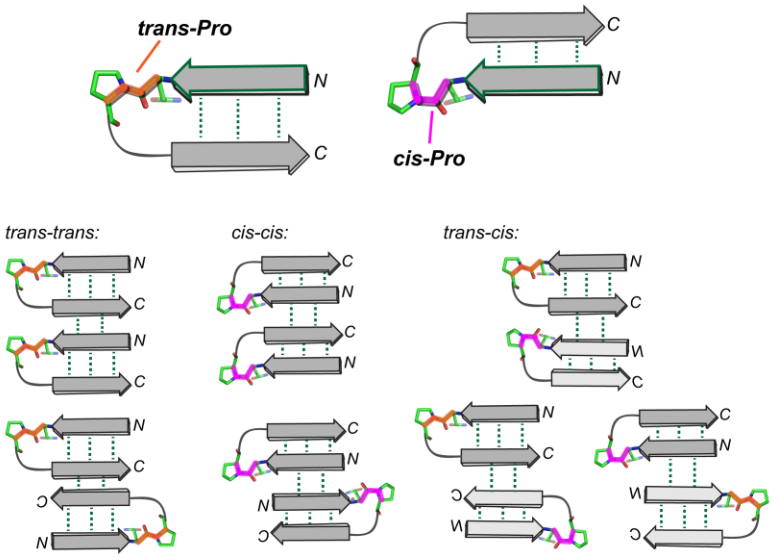

Figure 3.14 Proline isomers in turn structure.

Proline has two isomers, trans-proline (orange) and cis-proline (magenta). The Pro and preceding Gln in the turn are shown as sticks. For β-strands leading in the same direction, there are two possible ways to fold, 'down' (requiring trans-Pro) or 'up' (requiring cis-Pro). For anti-parallel arrangements of β-hairpins, turns incorporating all trans-Pro (trans-trans), all cis-Pro (cis-cis), or a mixture of trans- and cis-Pro (trans-cis) are available. Green dashed lines indicate backbone hydrogen bonding along the fibril axis. Both isomers are necessary for all turn structures to be possible.

Another feature of the PG turns is their heterogeneous nature, with both cis- and trans-Pro conformations. **Figure 3.14** illustrates how both the cis- and trans-Pro configurations at the end of a β-strand result in β-turn configurations that inherently point in opposite directions. In other words, the near-equal populations of β-II and β-IV turns indicate that there is an approximately random generation of β-hairpins in either of these conformations within the fibril sample. An intriguing question is whether this might extend to the hairpin structures that may be implicated in the original nucleation event. Above we have observed that antiparallel hydrogen bonded polyQ β-strands must feature complementary configurations (i.e. the two sets of NMR

signals). Thus, if it is the case that β-hairpin formation plays a key role early in the aggregation process, then it may be that there is an apparently random formation of the two kinds of β-hairpins. Independent of the above, once one of the "arms" of the hairpin forms a certain β-strand-like structure, this would then dictate that the opposite strand features the complementary conformation (as discussed above). Thus, the formation of the initial aggregation-prone polyQ conformation is a stochastic process, on multiple levels, which can result in multiple distinct conformations. In the polyQ context, these distinct conformations likely have very similar energetic and kinetic characteristics. The implication of these considerations is that any segment of the polyQ domain more-or-less randomly ends up in β-strands with two distinct conformations. Once an initial assembly is formed, also the templated growth can proceed in a random fashion, such that to resulting fibrils are composite assemblies in which each individual Gln is equally likely to exist in each of the different types of β-strands. This seemingly provides a molecular rationale for the observed MAS NMR features of polyQ aggregates, and is consistent with the body of evidence indicative of an important role for β-hairpin structures in polyQ misfolding and aggregation. An intriguing question is whether the lack of intermediate species in polyQ aggregation might imply that the structure of the early nuclei forming from the soluble polyQ ensemble would closely resemble the conformation of the peptides that end up in the amyloid-like fibril.

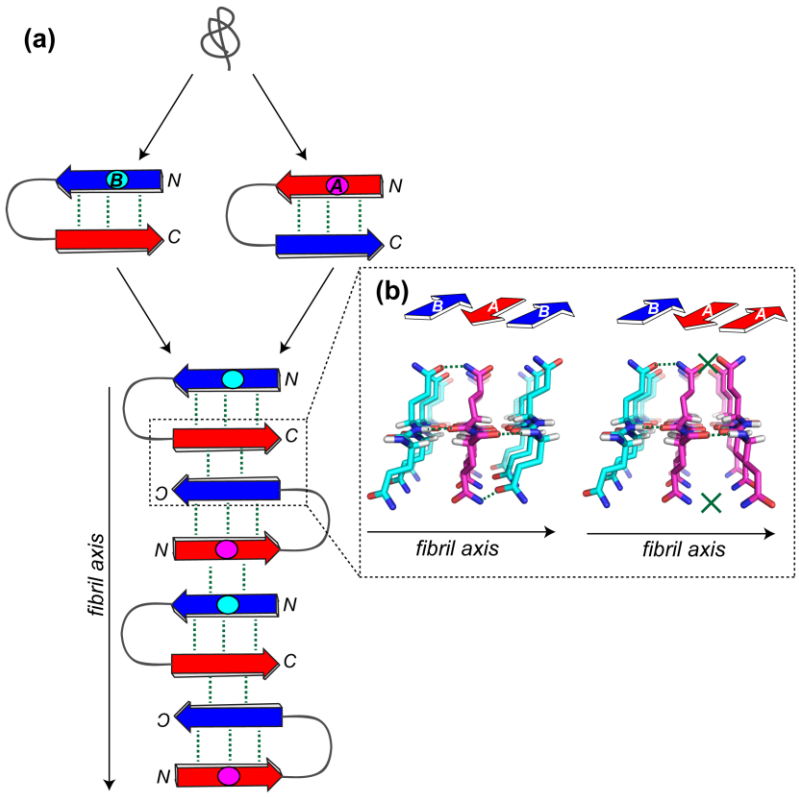

Figure 3.15 Assembly of Gln conformers 'a' and 'b' in polyQ fibrils.

(a) polyQ β-strands consist of Glns primarily in conformer 'a' (red) or conformer 'b' (blue). An individual Gln, indicated by the labeled circle has a 50/50 chance of taking on either of these conformations within an amyloid β-strand. With elongation along the fibril axis, β-strands must align in an anti-parallel fashion. In this model, the torsion angles of the conformer 'a' Gln side chains are assumed to have a χ_1 angle of -65° and a χ_2 angle of 180° and the conformer 'b' Gln side chains are assumed to have a χ_1 angle of 55° and a χ_2 angle of 180°. Due to the orientation of side chains, a β-strand of one conformer can only align with a strand of the opposite conformer, resulting in a stacking of alternating 'a' and 'b' β-strands along the fibril axis. Green dashed lines denote hydrogen bonding. (b) shows that Glns of opposite conformation must align with each other along the fibril axis in order for effective side chain hydrogen bonding to occur within an anti-parallel assembly. The right schematic shows the improper alignment of side chains when two conformer 'a' Glns are aligned along the fibril axis.

3.4.7 Applicability to Uninterrupted PolyQ and Disease-Relevant Proteins

Our data provide experimental support for the widely-made assumption that the PG-type insertions act by stabilizing β-hairpin structures in polyQ. One obvious and important question is whether or how this relates to the molecular features of uninterrupted polyQ and polyQ-containing proteins. The balance of evidence from aggregation kinetics [24, 25] along with structural data from FTIR [160] seems to support at least a likely role for β-hairpins at some stage of the misfolding and assembly process, but the degree to which there are β-hairpins in the mature fibrils remains a topic of debate.

3.5 CONCLUSION

Using a variety of MAS ssNMR measurements for structural constraints and relaxation parameters, we have gained new structural insight into the origin of the "signature motif" in the polyQ amyloid core. We have measured side chain χ_1 and χ_2 angles of the distinct Gln conformers, unexplainable by previously described models. We have thus proposed a new model for how these conformers are arranged within and between β-strands in the amyloid core. Insertion of Pro-Gly motifs in the middle of a Gln stretch gave insight into turn conformations connecting β-strands in the amyloid core. We have observed how these Pro residues are incorporated into the polyQ amyloid structure and the localized disruption by Pro-Gly insertions. These results rationalize the unusual ability of polyQ amyloid to accommodate the presence of Pro-Gly mutations. Our data complement earlier mechanistic studies [25] and provide additional support for models of polyQ aggregation involving β-hairpin structures. A more detailed characterization of the structural features of these peptides and polyQ amyloid in general, is ongoing.

3.6 ACKNOWLEDGEMENTS

The authors thank Mike Delk for technical assistance, and Seth Horne for helpful discussions. We acknowledge funding support from the University of Pittsburgh and the National Institutes of Health (grants R01 AG019322 to R.W. and P.v.d.W and R01 GM099718 to R.W.).

4.0 STRUCTURAL CHARACTERIZATION OF THE CAVEOLIN SCAFFOLDING DOMAIN IN ASSOCIATION WITH CHOLESTEROL-RICH MEMBRANES

Adapted with permission from: **Hoop, C.L.**, Sivanandam, V.N., Kodali, R., Srnec, M.N., and van der Wel, P.C.A. Structural Characterization of the Caveolin Scaffolding Domain in Association with Cholesterol-Rich Membranes, *Biochemistry 51*, 90-99. Copyright (2012) American Chemical Society.

Author contributions: C.L.H. and V.N.S. prepared samples and performed and analyzed ssNMR and CD experiments. C.L.H. analyzed computational secondary structure predictions. R.K. performed FTIR measurements. M.N.S. analyzed secondary structures within previously reported CRAC motifs. C.L.H., V.N.S., R.K., M.N.S., and P.v.d.W. wrote the publication.

Members of the caveolin protein family are implicated in the formation of caveolae, play important roles in a number of signaling pathways and in the regulation of various proteins. We employ complementary spectroscopic methods to study the structure of the caveolin scaffolding domain (CSD) in caveolin-1 fragments, while bound to cholesterol-rich membranes. This key domain is thought to be involved in multiple critical functions that include protein recognition, oligomerization, and cholesterol binding. In our membrane-bound peptides, residues within the flanking intramembrane domain (IMD) are found to adopt an α-helical structure, consistent with

its commonly believed helical hairpin conformation. Intriguingly, in these same peptides, we observe a β-stranded conformation for residues in the CSD, contrasting with earlier reports, which commonly do not reflect β-structure. Our experimental data based on solid-state NMR, CD, and FTIR are found to be consistent with computational analyses of the secondary structure preference of the primary sequence. We discuss how our structural data of membrane binding Cav fragments may match certain general features of cholesterol-binding domains and could be consistent with the role for CSD in protein recognition and homo-oligomerization.

4.1 INTRODUCTION

Caveolae are flask-shaped and cholesterol-rich invaginations in plasma membranes found to participate in many cellular functions across various cell types [29, 197, 198]. Their roles include lipid uptake and regulation, molecular transport, cell adhesion, and signal transduction [28, 29]. The caveolin family of proteins is thought to play a central role in the formation and functioning of these lipid raft-like domains. There are three isoforms, caveolin-1 (Cav1), caveolin-2, and caveolin-3, of which Cav1 is most common. Found in a wide array of mammalian cells, Cav1 is implicated in multiple diseases, including cancers and muscular disease [27, 199]. It contains cytoplasmic N- and C-terminal domains that sandwich the membrane-association domains: the caveolin scaffolding domain (CSD; residues 82-101) and the intramembrane domain (IMD) spanning roughly residues 102-134 (**Figure 4.1**) [197]. The latter is thought to form a unique α-helical hairpin that does not completely traverse the membrane

[27, 200, 201]. These domains are assisted in their membrane-binding roles by three palmitoylated cysteines, which are not essential for localization to caveolae [202-204], but do appear critical for other functionality, such as the transportation of cholesterol [204].

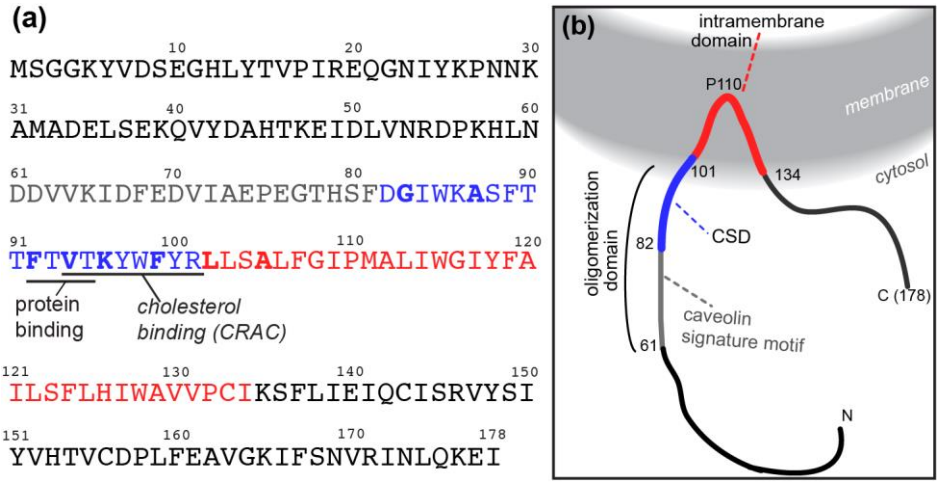

Figure 4.1 Primary sequence and schematic illustration of caveolin-1.

Key domains are indicated: the CSD in blue, containing the cholesterol-binding CRAC motif and the residues required for binding of protein partners (underlined). Bold residues indicate sites U-^{13}C,^{15}N-labeled in one or more peptides.

The most prominent domain is the CSD, as residues within it are not only responsible for membrane binding, but are also critical for oligomerization, protein interactions, and cholesterol recognition [31, 205]. Residues within the CSD are required for oligomerization of Cav monomers into homo-oligomers of 14-16 proteins, which themselves assemble into higher-order oligomers during the formation of caveolae [205]. On top of this, the $F_{92}TVT_{95}$ segment within the CSD is important for signaling, as it is required for interaction with binding protein partners (e.g. G-protein α-subunits) [33, 205, 206]. This binding event involves a consensus motif in the partner protein with high aromatic content, sometimes referred to as a caveolin-binding motif (CBM). A similar motif is also found in caveolin itself, raising the possibility of structural

similarities between protein partner binding and Cav homo-oligomerization [33]. Finally, the CSD is ascribed a role in the recognition and binding of cholesterol, which is highly concentrated in caveolae and needed for their formation. More precisely, formation of caveolae strictly requires tight binding of Cav to cholesterol [207]. This functionality is thought to localize to a cholesterol recognition/interaction amino acid consensus (CRAC) motif [28] in residues V_{94}TKYWFYR$_{101}$ [32]. Thus, despite its short length, this 20-residue segment appears to incorporate an array of critical functionalities in overlapping sequence elements.

Unfortunately, its molecular mechanism of action and structural features remain unclear. Many reports indicate significant but varying degrees of α-helical structure in the CSD. A number of experimental studies have reported partially or even extensively α-helical CSD [32, 208]. This is in part supported by computational analysis of the primary sequence [27, 208-211]. A helical conformation has been employed to rationalize Cav1 binding partner recognition via the CSD [212], as well as the cholesterol-binding of CRAC motifs in general [213]. Comparisons of the various experimental studies are complicated by differing sample conditions and Cav1 constructs, some of which are rather short fragments [32, 211]. It has been suggested that use of overly short Cav1 fragments may have caused a loss of helicity due to peptides being truncated within secondary structure elements [208]. Experimental conditions also often involve vesicular or micellar samples that lack cholesterol and may not emulate the native, highly cholesterol-rich membrane environment of the caveolae [214]. In contrast to most other studies, an integrative bioinformatics study by Spisni et al suggested an antiparallel β-hairpin for residues 84-94, based on their computational analysis and previous functional and mutational results [215].

To explore some of these ambiguities regarding the structure of Cav1, particularly the CSD, we here present results from our studies of Cav1 fragments in cholesterol-rich lipid bilayer

samples. Through the use of ssNMR, we examine the structure of the peptide fragments while bound to the membrane, and also observe their modulation of the membrane structure and dynamics. Solid-state NMR is a powerful tool for membrane protein studies, as it is compatible with a wide range of sample compositions and allows characterization of both lipids and proteins or peptides [216-218]. This applies to both MAS NMR measurements as well as 'static' experiments on pelleted or aligned membranes. Indeed, previous work on shorter CSD fragments employed ssNMR for the characterization of their effect on the membrane [32]. We perform static and MAS ssNMR experiments, supported with circular dichroism (CD) and Fourier transform infrared (FTIR) spectroscopy to probe selected Cav1 fragments. These include a short peptide that reflects the isolated CSD (Cav_{83-102}) and a longer construct (Cav_{82-109}), which incorporates part of the IMD. Previous reports suggested that the former has a mixed secondary structure, but that the latter has a predominantly α-helical conformation, at least in dodecylphosphocholine (DPC) micelles [32, 208]. Here, we examine their structures in cholesterol-rich lipid bilayers designed to more closely mimic caveolae, which should provide a better context to probe the cholesterol-binding and oligomerization characteristics of these peptides. Our CD and FTIR data indicate an increased α-helicity in the longer peptide (in agreement with some earlier reports), but ssNMR reveals that this α-helicity is specifically located in the putative first helix of the IMD and that residues within the CSD segment adopt a β-sheet conformation even in the presence of the flanking IMD. This turns out to be consistent with structural propensities of these fragments, as well as full-length Cav1, as indicated by various primary-sequence-based secondary structure prediction algorithms. We conclude by discussing our observations in the light of previous reports, along with possible implications for the structure and function of caveolin and its multifaceted scaffolding domain.

4.2 EXPERIMENTAL PROCEDURES

4.2.1 Sample Preparation

Appropriately Fmoc- and side-chain-protected ^{13}C, ^{15}N-labeled amino acids were purchased from Cambridge Isotope Laboratories (Andover, MA) and Isotec (Sigma-Aldrich, St. Louis, MO). Site-specifically ^{13}C, ^{15}N-labeled and unlabeled peptides (**Table 4.1**) were synthesized by Fmoc solid-phase chemistry and purified via reversed phase HPLC to greater than 95% purity by the W.M. Keck Facility at Yale University and New England Peptides (Gardner, MA). 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) was obtained from Avanti Polar Lipids (Alabaster, AL) and cholesterol was obtained from Sigma-Aldrich (St. Louis, MO); both were used without further purification. Membrane samples containing a 1:1 molar ratio of phospholipid/cholesterol were prepared following modifications of previously published protocols [32, 219]. Peptide dissolved in trifluoroethanol (TFE) was combined with co-dissolved phospholipid and cholesterol in chloroform/methanol (2/1, v/v). A dry peptide/lipid/cholesterol film, obtained by drying under a N_2-stream and overnight exposure to high vacuum, was resuspended in excess HEPES buffer (20mM HEPES with 1mM EDTA and 150mM NaCl at pH 7.0), sonicated for 10-15 min, and subjected to several cycles of freeze-thawing. Negatively-stained transmission electron microscopy and dynamic light scattering measurements (not shown) indicated the resulting vesicles were typically ~100 nm in diameter.

Table 4.1 Sequences, labeling schemes, and amounts of isotopically labeled peptides used for the MAS ssNMR samples (in 1:1 POPC/cholesterol).

Labeled residues are underlined in the primary sequence.

ID	Labeling	Sequence	Peptide amount (mg)
p1	U-^{13}C,^{15}N-[A_{87}, F_{92}, V_{94}, K_{96}]	DGIWK<u>A</u>SFTT<u>F</u>T<u>V</u>T<u>K</u>YWFYRLLSALFGI (Cav_{82-109})	3.7
p2	U-^{13}C,^{15}N-[G_{83}, V_{94}, F_{99}, L_{102}, A_{105}]	D<u>G</u>IWKASFTTFT<u>V</u>TKYW<u>F</u>YR<u>L</u>LS<u>A</u>LFGI (Cav_{82-109})	2.65
p3	U-^{13}C,^{15}N-[G_{83}, A_{87}]	EGTHSFD<u>G</u>IWK<u>A</u>SFTTFTVTKYWFYRLLSALFGI (Cav_{76-109})	3.6

4.2.2 Solid-State NMR

Membrane samples, with or without 10 mol% peptide, were pelleted into 4mm or 3.2mm zirconia MAS rotors (Bruker Biospin, Billerica, MA) by centrifugation and were kept fully hydrated and unfrozen at all times. All experiments were carried out on a wide-bore Bruker Avance I spectrometer operating at 600 MHz ^{1}H Larmor frequency (14.3T) using either a Bruker wide-bore 4mm CPMAS triple-channel or standard-bore 3.2mm CPMAS EFree HCN probe. The MAS experiments employed a spinning rate $\omega_r/2\pi$ = 8 kHz, while cooling the sample using pre-cooled gas, at all times maintaining the sample temperature above the phase transition temperature of the lipids. Static and MAS ^{31}P spectra were acquired with a wide-bore 4mm double-channel (^{1}H,^{31}P) CP MAS probe (Bruker) in absence and presence of sample spinning, applying 50kHz two-pulse phase modulation (TPPM) ^{1}H decoupling [220]. Line shape analysis and integration were performed using Topspin software (Bruker Biospin, Billerica, MA). ^{31}P chemical shifts were referenced to phosphoric acid via external referencing to hydroxyapatite

[221]. Assignments of peptide resonances were primarily carried out through analysis of 2D ^{13}C-^{13}C experiments employing ^{1}H-^{13}C cross polarization (CP) followed by 25-100ms DARR mixing [35] with 83kHz TPPM decoupling during acquisition and evolution. Long distance contacts were observed via 2D ^{13}C-^{13}C experiments utilizing 400ms PDSD mixing with 71-83 kHz TPPM decoupling during acquisition and evolution. Assignments of the POPC and cholesterol resonances were done with reference to previously published literature and are supported by 2D ^{1}H-^{13}C HETCOR experiments (not shown) [222, 223]. Spectra were processed and analyzed using the Bruker Topspin, NMRPipe [171], Sparky [224], and CCPNMR/Analysis programs [225]. ^{13}C chemical shifts were referenced to dilute aqueous DSS via external referencing to adamantane, and ^{15}N chemical shifts were indirectly referenced to liquid ammonia [226]. Additional experimental details are available in **Table A.5** in **Appendix C**.

4.2.3 FTIR and CD Spectroscopy

For FTIR, samples containing 5mol% or 10mol% $Cav_{82\text{-}109}$ or $Cav_{83\text{-}102}$ in 1:1 POPC/cholesterol or DPC micelles were obtained as pellets via centrifugation. I acknowledge Dr. Ravindra Kodali for performing the FTIR experiments. The pellet was placed between two CaF_2 windows on an MB series spectrophotometer (ABB Bomem, Quebec City, QC, Canada) and analyzed with PROTA software (Biotools Inc., Jupiter, FL). Spectra were recorded at 4 cm^{-1} resolution and at room temperature (averaging over 400 scans). Buffer correction was applied by subtracting the buffer spectrum interactively until a flat baseline was obtained between 1700 and 1800 cm^{-1}. Peak volumes were estimated by fitting with Gaussian line shapes using the PeakFit routine in SigmaPlot (Systat Software, San Jose, CA).

Cav_{82-109} and Cav_{83-102} containing samples for CD were prepared in either POPC vesicles (1:100 peptide-lipid ratio) or DPC micelles (4 mM) in a 13.3 mM phosphate buffer at pH ~6.25, resulting in final peptide concentrations of 10 – 25 μM. CD spectra were recorded at 25°C using a Jasco J-810 spectropolarimeter with a 1mm path length quartz cuvette (Jasco Inc., Easton, MD), averaging over 10 scans with a scanning speed of 50nm/min in steps of 1 nm.

4.2.4 Sequence Analysis and Secondary Structure Prediction

The secondary structure of full-length Cav1 and fragments Cav_{82-109} and Cav_{83-102} was predicted via selected algorithms. The PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/), Proteus2 (http://wishart.biology.ualberta.ca/proteus2), and PredictProtein (http://www.predictprotein.org/) algorithms were accessed via their respective websites [227-229].

4.3 RESULTS

4.3.1 Peptide Incorporation into Model Membranes Using Static and MAS ssNMR

In order to confirm the incorporation of the peptides into the lipid membranes, and to probe their localization within the membrane, we applied a number of ssNMR experiments. In analogy with previous work [32], we examined the peptides' effect on the lipids through basic ^{13}C and ^{1}H MAS NMR experiments. In our hands, the POPC and cholesterol chemical shift changes upon peptide introduction were relatively small and did not unequivocally identify more significant effects on particular parts of the lipids. In various independently prepared samples, introduction of CSD peptides does consistently cause a broadening of the carbon line-widths, likely due to a decrease in membrane fluidity. **Figure 4.2a-c** shows the natural abundance ^{13}C signals from 1:1 POPC/cholesterol in presence of 10mol% Cav_{82-109} **(a)**, with 10mol% Cav_{83-102} **(b)** or without peptide **(c)**. It is notable that a single signal is observed for each lipid and cholesterol site, indicating a single homogeneously structured population of the lipids.

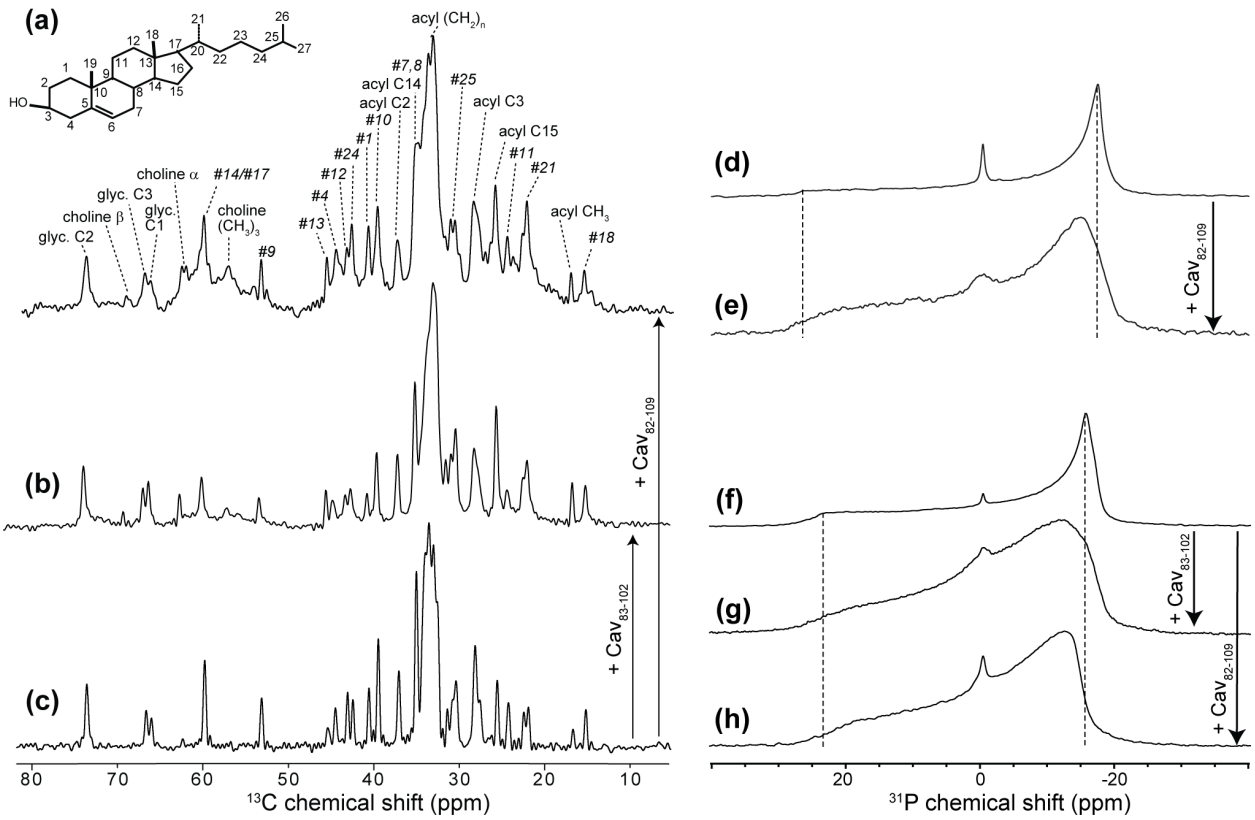

Figure 4.2 SSNMR shows effects of Cav1 fragment peptides on the membranes.

^{1}H-^{13}C CPMAS spectra of 1:1 POPC/cholesterol in presence of (a) 10mol% Cav_{82-109} (unlabeled), (b) 10 mol% Cav_{83-102} (unlabeled) or (c) without peptide. Selected assignments of POPC and cholesterol are shown, the latter as italicized numbers (#; see insert). In presence of the peptides, ^{13}C signals throughout the lipids are broadened and some previously missing mobile POPC signals become visible. Static ^{31}P NMR of POPC without (d) or with (e) 10 mol% Cav_{82-109}, and 1:1 POPC/cholesterol in absence of peptide (f) or in presence of (g) 10 mol% Cav_{83-102} or (h) 10mol% Cav_{82-109} shows that the lipids predominantly form a bilayer, with less than 5% present as isotropic signals. Peptide incorporation changes the lipid headgroup dynamics, seen as a reduction of the bilayer ^{31}P CSA (see dashed vertical lines). Measurements were done with full hydration, at 300 K and 600MHz ^{1}H frequency.

To probe the behavior of the phospholipid head groups, we applied static ^{31}P NMR experiments that are informative of the lipids' macroscopic assembly into distinct lipid phases and the dynamics of individual lipids. Static 1D ^{31}P NMR spectra of 1:1 POPC/cholesterol

mixtures are dominated by broad signals with a lineshape typical of uniaxial motional averaging of the ^{31}P chemical shift anisotropy (CSA) (**Figure 4.2f**) as part of the complex motions that characterize phospholipid bilayers [230]. The apparent uniaxial rotation reflects rapid Brownian motion of the phospholipids within the fluid bilayer. In addition to the bilayer phase, certain samples also feature a small isotropic peak (close to 0 ppm), constituting at most 5% of the overall intensity (as determined by line shape simulations). Such isotropic ^{31}P NMR signals indicate that a small portion of the phospholipid head groups undergoes rapid isotropic motional averaging, as is found in small vesicles, micelles, or highly curved non-bilayer phases [231, 232]. Note that these isotropic signals are not specifically correlated to the presence of peptide but are also seen in lipid-only samples. Compared to POPC bilayers alone (**Figure 4.2d**), introduction of high levels of cholesterol (**Figure 4.2f**) leads to a notable decrease of the ^{31}P CSA [233]. This is most likely due to an increased wobbling motion of the PC head group [230], which can be rationalized based on the fact that cholesterol mostly occupies the hydrophobic core of the membrane, lacks a large head group, and thus allows for more space and motion of the PC head group [234].

In the presence of the Cav1 fragments, the PC head groups still show a dominant bilayer line shape (**Figure 4.2g-h**). The Cav1 peptides do significantly modulate the observed inhomogeneous and homogeneous line broadening of the lipid ^{31}P signals. They have a concentration-dependent (**Figure 4.3**) narrowing effect on the ^{31}P CSA. In analogy to cholesterol's effect, this may indicate incorporation into the membrane's hydrophobic core with less interaction with the lipid head groups. This decrease in the inhomogeneous broadening is more pronounced for the $Cav_{82\text{-}109}$ peptide (**Figure 4.2h**) than the isolated CSD (**Figure 4.2g**). Addition of the peptide also causes an *increase* in the homogeneous linewidth (**Figure 4.2e, g,**

and h). Such effects have previously been seen with other membrane-binding compounds, and are most likely due to an interference with the overall motion of the phospholipids [235]. In POPC (**Figure 4.2e**) and, to a lesser extent, POPC/cholesterol bilayers (**Figure 4.2g-h**), the rapid Brownian motion of each lipid in the fluid bilayer helps to reduce the ^{31}P-^{1}H dipolar interactions, which can otherwise cause an "intrinsic broadening" of the ^{31}P signal of individual lipids [235, 236]. The incorporation of the Cav1 fragments interferes with the rapid lipid motion and thus results in an increased homogeneous broadening of the ^{31}P NMR line shape.

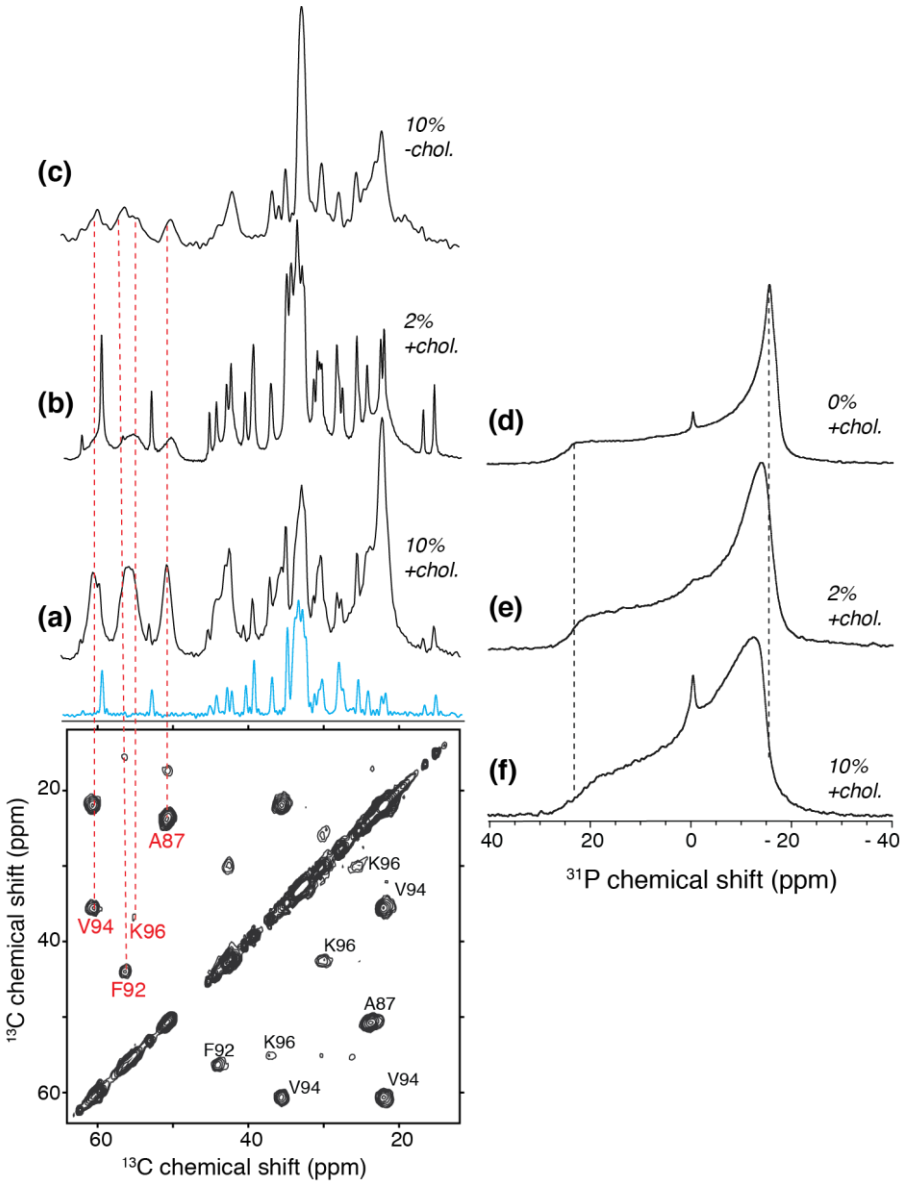

Figure 4.3 $Cav_{82\text{-}109}$ concentration effects.

(a-c) Comparison of ^{13}C 1D spectra: (a) 10mol% $Cav_{82\text{-}109}$ in 1:1 POPC/cholesterol, (b) 2 mol% $Cav_{82\text{-}109}$ in 1:1 POPC/cholesterol, (c) 10 mol% $Cav_{82\text{-}109}$ in pure POPC. A 1:1 POPC/cholesterol ^{13}C spectrum is provided in cyan for reference of lipid and cholesterol signals (ref. ***Figure 4.2*** *for specific assignments). The ^{13}C-^{13}C DARR 2D (from* ***Figure 4.6****) shows the peptide assignments (10 mol% in POPC/cholesterol). The most visible Cα signals are marked in red. (d-f) Comparison of ^{31}P lineshapes with increasing concentrations of $Cav_{82\text{-}109}$. (d) 1:1 POPC/cholesterol in absence of peptide, (e) 2 mol% $Cav_{82\text{-}109}$ in 1:1 POPC/cholesterol, and (f) 10 mol% $Cav_{82\text{-}109}$ in 1:1 POPC/cholesterol. A concentration-dependent narrowing of the ^{31}P CSA is observed upon addition of peptide.*

The above observations indicate an intimate interaction of these peptides with the lipid membranes, without signs of peptide-induced heterogeneity. The peptides seem to partition preferentially into the hydrophobic part of the bilayer, providing more space for POPC head group motion, while reducing the overall fluidity of the membrane.

4.3.2 Peptide Secondary Structure Content

Having probed the peptides' incorporation into the mixed membranes, we also characterized the secondary structure of the membrane-associated peptides. FTIR studies were performed on $Cav_{82\text{-}109}$ and $Cav_{83\text{-}102}$ peptides in a 1:1 POPC/cholesterol mixture, prepared identically to the NMR samples (**Figure 4.4**). $Cav_{82\text{-}109}$ features C=O vibrations characteristic for both α-helix (1654 cm^{-1}) and β-strand (1623 and 1693 cm^{-1}) [237]. Based on peak deconvolution (see **Figure 4.4**), relative amounts of approximately 60% α-helical and 40% β-sheet are estimated for $Cav_{82\text{-}109}$ (**Table 4.2**). In contrast, the shorter $Cav_{83\text{-}102}$ displays little α-helical content (**Figure 4.4b**). The phospholipid's ester C=O vibration is also observed at 1730 cm^{-1} [238].

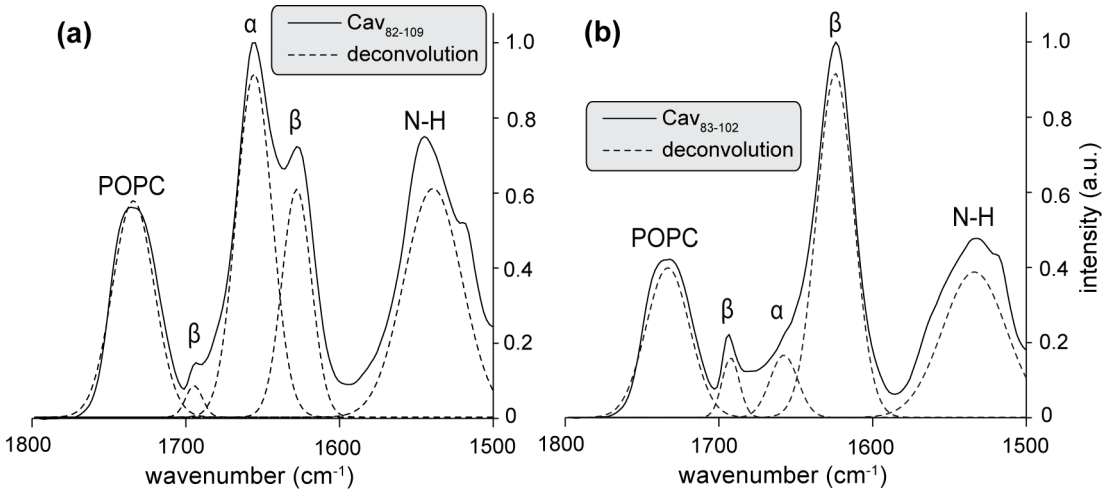

Figure 4.4 Quantifying secondary structure in Cav1 fragment peptides in a cholesterol-rich lipid bilayers.

FTIR spectra of (a) 10 mol% Cav_{82-109} and (b) 10 mol% Cav_{83-102} in 1:1 POPC/cholesterol. Deconvolutions for quantifying secondary structure content are shown as dashed for their respective data sets. Results are consistent with mixed secondary structure content in the longer peptide, and predominantly β-sheet structure in the shorter construct.

Table 4.2 Secondary structure content estimates from FTIR.

The relative amounts of α-helix and β-sheet content of both peptides (Cav_{82-109} and Cav_{83-102}) in POPC/cholesterol were estimated by deconvolution of the FTIR spectra as shown in ***Figure 4.4*** *and* ***Figure*** **4.5**.

Sample	α-helix (%)	β-sheet (%)
10mol% Cav_{82-109}	63	37
10mol% Cav_{83-102}	11	89
5mol% Cav_{82-109}	56	44
5mol% Cav_{83-102}	15	85

The FTIR results thus indicate a notable increase in α-helical content in the longer peptide, an observation that is consistent with the work by Le Lan et al, employing CD spectroscopy [208].

These earlier studies were done in DPC, rather than cholesterol/POPC membranes. We obtained similar FTIR results in DPC micelles (not shown), and CD spectra in POPC that indicated a low level of helicity in the shorter peptides and an increased α-helical content for the longer peptide (**Figure 4.5**).

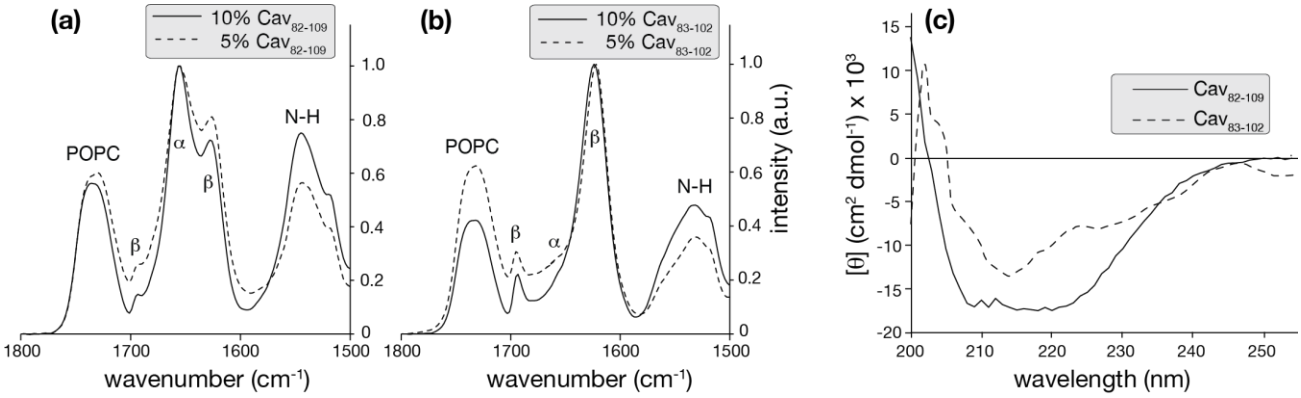

Figure 4.5 Structural elements in Cav1 fragment peptides from FTIR and CD data.

FTIR spectra for (a) $Cav_{82\text{-}109}$ and (b) $Cav_{83\text{-}102}$ in 1:1 POPC/cholesterol. For each peptide, the resulting data with 10mol% peptide (solid) and 5mol% peptide (dashed) is shown. Relative amounts of α-helix and β-strand structure have been estimated by peak deconvolution and are tabulated in ***Table 4.2, above****. (c) CD spectra of 25uM $Cav_{82\text{-}109}$ (solid) and 10uM $Cav_{83\text{-}102}$ (dashed) in POPC at a 1:100 peptide-to-lipid ratio at 25ºC. Although poor spectral quality due to scattering prevented a reliable quantitative analysis, it is clear there is an increase in helical structure in the longer peptide as previously reported in DPC by Le Lan et al. [208].*

Thus, both CD and FTIR indicate an increase in α-helical content in $Cav_{82\text{-}109}$, but it appears that a substantial fraction remains non-α-helical. These data do not indicate the localization of the helix and could be explained by sample heterogeneity.

4.3.3 Site-Specific Structure Analysis by MAS ssNMR

To site-specifically probe the mixed secondary structures of $Cav_{82\text{-}109}$, MAS ssNMR was applied to partly $^{13}C,^{15}N$-labeled peptides. Two $Cav_{82\text{-}109}$ peptides with $^{13}C,^{15}N$-labeling of residues in both the CSD and IMD were incorporated into 1:1 POPC/cholesterol membranes. The first peptide (p1) featured U-$^{13}C,^{15}N$-labeled residues in sites Ala87, Phe92, Val94, and Lys96; and the second (p2) was labeled in residues Gly83, Val94, Phe99, Leu102, and Ala105 (**Table 4.1**). Probing residues Phe92 and Val94 yields structural information on the FTVT protein binding segment [206], while analogously measuringVal94, Lys96, and Phe99 provides insight on the CRAC motif [32]. Labeling of Val94 in both peptides permitted verification of unchanged Val94 shifts between different samples, which was indeed the case (**Figure 4.6**). Chemical shift assignments were obtained using 2D ^{13}C-^{13}C DARR [35] experiments (e.g. **Figure 4.7**), supplemented with 2D ^{15}N-^{13}C correlations (not shown), and are available in **Table A.6** in **Appendix C**. In most cases single resonances were observed, indicating the presence of a single well-defined conformation, although motional averaging that is fast on the NMR time scale cannot be excluded. We do observe weaker additional resonances for Ala87 and Val94 (in both peptides), indicating a less-populated secondary conformation of those residues. Reduced intensities of certain peaks (e.g. the labeled Phe side chains) are consistent with conformational exchange of the amino acid side chains.

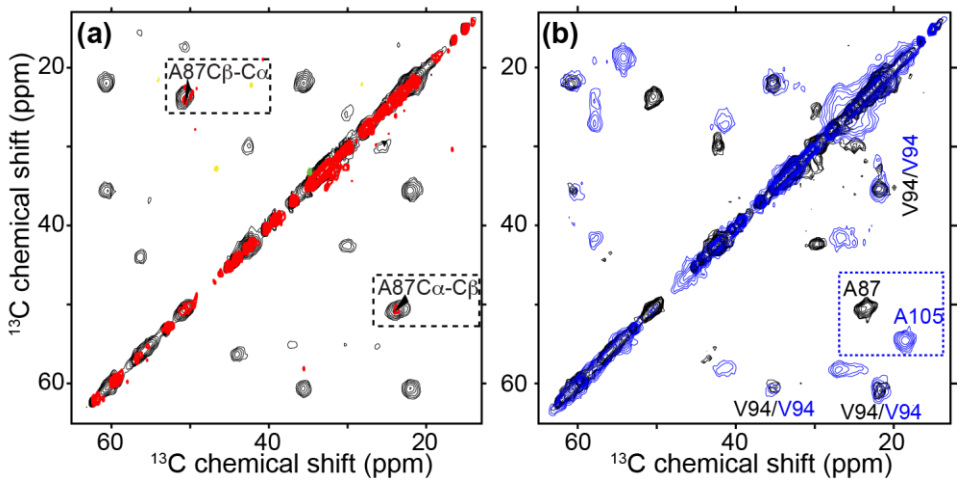

Figure 4.6 Overlay of 2D ^{13}C-^{13}C ssNMR spectra of differently labeled Cav1 fragment peptides samples.

Shown is the intra-aliphatic region of spectra with 25 ms DARR mixing, providing 1-2 bond transfers. Panel (a) compares the 28- and 34-residue long peptides p1 (black) and p3 (red), where A87 in the CSD is uniformly ^{13}C-labeled in both. The Cα and Cβ chemical shifts are unchanged and are indicative of β-sheet secondary structure. In (b), we compare peptides p1 (black) and p2 (blue), showing that the V94 shifts are reproducible, but the two Ala residues in the CSD (A87) and intramembrane domain (A105) have very different shifts due to the difference in secondary structure. Spectra were obtained at 600 MHz ^{1}H frequency, 8 kHz MAS and at 283 K.

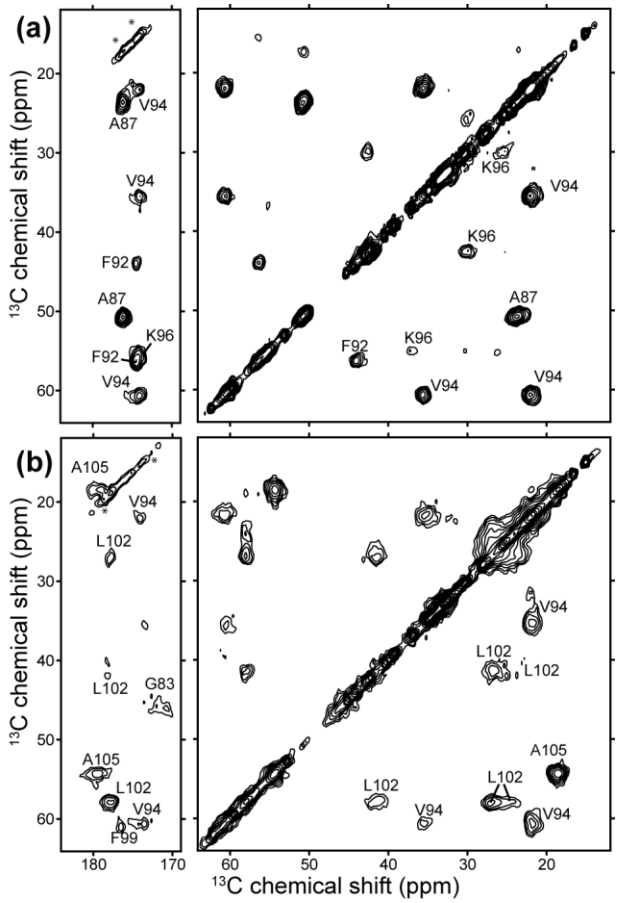

Figure 4.7 ^{13}C Chemical shift assignments for isotopically labeled residues in Cav1 fragment peptides.

2D ^{13}C-^{13}C ssNMR spectra using 25 ms DARR mixing, providing 1-2 bond transfers. Measurement on samples (a) p1 and (b) p2 are shown; both obtained at 600 MHz ^{1}H frequency and 8 kHz MAS at 283 K. Aliphatic-to-carbonyl (left) as well as intra-aliphatic (right) spectral regions are shown.

We examined the residue-specific secondary structure by comparing our chemical shifts to the corresponding shifts of known protein structures, in an approach known as chemical shift indexing (CSI) [239, 240]. The difference of ^{13}C chemical shifts from random coil shifts [50] correlates in a predictable fashion to the local secondary structure. As shown in **Figure 4.8**, these data indicate that residues Ala87, Phe92, Val94, and Lys96, spanning the protein binding and CRAC motifs, adopt a β-sheet conformation, whereas residues Phe99 (part of the CRAC motif),

Leu102 and Ala105 (in the IMD) adopt an α-helical structure. Gly83, near the N-terminus of this peptide fragment, appears to lack a clearly defined secondary structure.

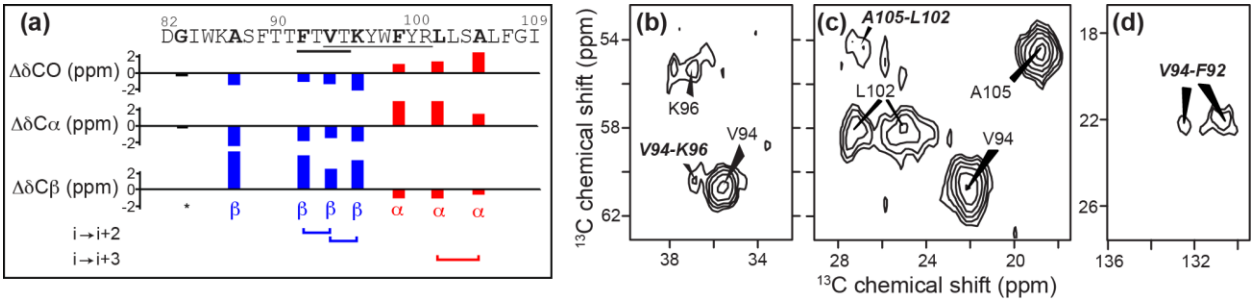

Figure 4.8 Secondary structure determination from ssNMR.

(a) SSNMR indicates α-helical conformation (red) for labeled sites in the IMD, whilst most CSD residues have a β-conformation (blue). Residues of the CRAC motif (V_{94} – Y_{100}) span the helix-strand boundary, and the protein binding motif (F_{92} – T_{95}) is in a β-strand. The CSI shows the difference (Δδ) of C', Cα, and Cβ chemical shifts from random coil values. Black bars lack a defined secondary structure (e.g. G83). Also indicated are the observations of i → i+2 contacts between F92-V94 and V94-K96 (b,d) and an i → i+3 contact between L102-A105 (c). The spectra are from long-distance ^{13}C-^{13}C experiments with 400 ms PDSD mixing, obtained at 600 MHz ^{1}H frequency and 8 kHz MAS at 283 K.

Given concerns about the effect of truncation of secondary structure elements [208], we also examined a longer peptide $Cav_{76\text{-}109}$ (p3), which in the native sequence is flanked on both sides by prolines (P75 and P110), which often act as secondary structure breakers. Their effect on membrane-spanning helices seems to be more subtle, yet they do tend to destabilize these α-helices [241]. We incorporated isotopic labels at Gly83 and Ala87 as probes of the CSD secondary structure, and found that these residues had the same chemical shifts (and thus structures) as in the $Cav_{82\text{-}109}$ fragment (**Figure 4.6**).

The β-sheet and α-helical structural elements in $Cav_{82\text{-}109}$ were further probed in 2D ^{13}C-^{13}C experiments with 400ms PDSD mixing. These conditions permit ^{13}C-^{13}C magnetization

transfer over longer distances and revealed not only intra-residue contacts within labeled residues, but also specific inter-residue contacts. As summarized in **Figure 4.8**, residue Val94 is observed to interact with both Phe92 and Lys96, consistent with i → i+2 contacts typical of β-sheet structure. Similarly, residue Leu102 interacts with residue Ala105, reflecting an i → i+3 contact characteristic of α-helical structure (and one that would not likely be observed in a β-strand). The observation of these i → i+2 and i → i+3 contacts are all consistent with the secondary structure elements identified from the CSI analysis above and further solidify the secondary structure assignments (**Figure 4.8**).

4.3.4 Structure Prediction

Our observations of a partially β-sheet structure localized within the CSD (specifically including the labeled residues between positions 87 and 96) seem to match results from a computational study that predicted an anti-parallel β-hairpin motif in residues 84-94 [215]. However, a number of other studies report a preference for α-helical structure from primary sequence analysis [27, 210]. To explore this in more detail, we have also applied a number of prediction algorithms on full-length Cav1 and selected fragments. The results of predictions using the PSIPRED algorithm [227] are shown in **Figure 4.9**, suggesting a preference for an extended (possibly β-sheet) conformation throughout the CSD segment, followed by an α-helical structure in the IMD, not only in full length Cav1 but also in the $Cav_{82\text{-}109}$ fragment (grey box in **Figure 4.9**). Moreover, consistent with our and earlier experimental results [208], truncation of the α-helical segment (i.e. ending at residue 102 rather than just before the native helix-breaker Pro110) is predicted to destabilize the helical conformation (e.g. in $Cav_{83\text{-}102}$). However, there

appears to be little indication of a fully α-helical CSD. Similar results were obtained with the PROTEUS2 and PredictProtein prediction algorithms [228, 229], both of which suggest mixtures of β- and α-structure within the CSD (see **Figure 4.10**). Overall, the results of all three algorithms are also somewhat similar to the consensus results presented by Spisni et al, but differ from the reports of more extensive helical preferences. Of course, it remains unseen how reliable any of these generally globular-protein-based algorithms will prove to be for this oligomerizing, cholesterol-binding and likely monotopic membrane protein.

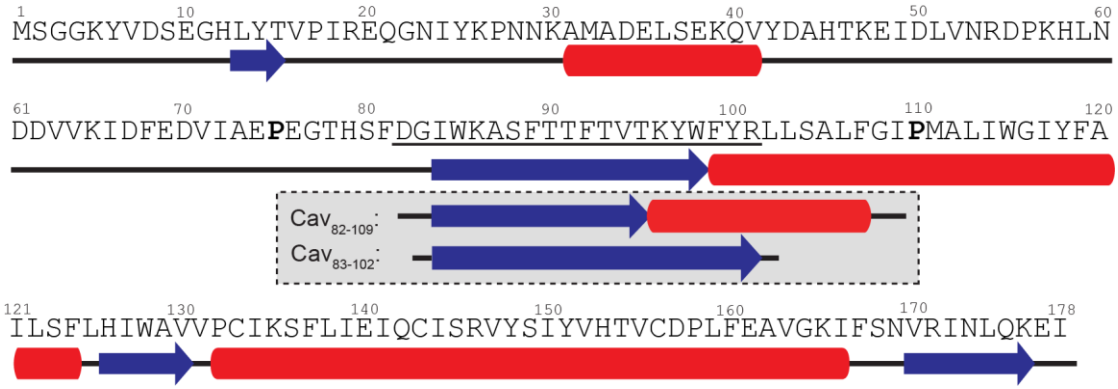

Figure 4.9 Secondary structure prediction of Cav1 and fragments.

PSIPRED analysis of full-length Cav1 (Cav_{1-178}) indicates a propensity for β-sheet structure in the CSD (underlined). Predictions for fragments Cav_{82-109} and Cav_{83-102} (grey insert) suggest a sensitivity to helix truncation, but always include a CSD with significant β-sheet structure.

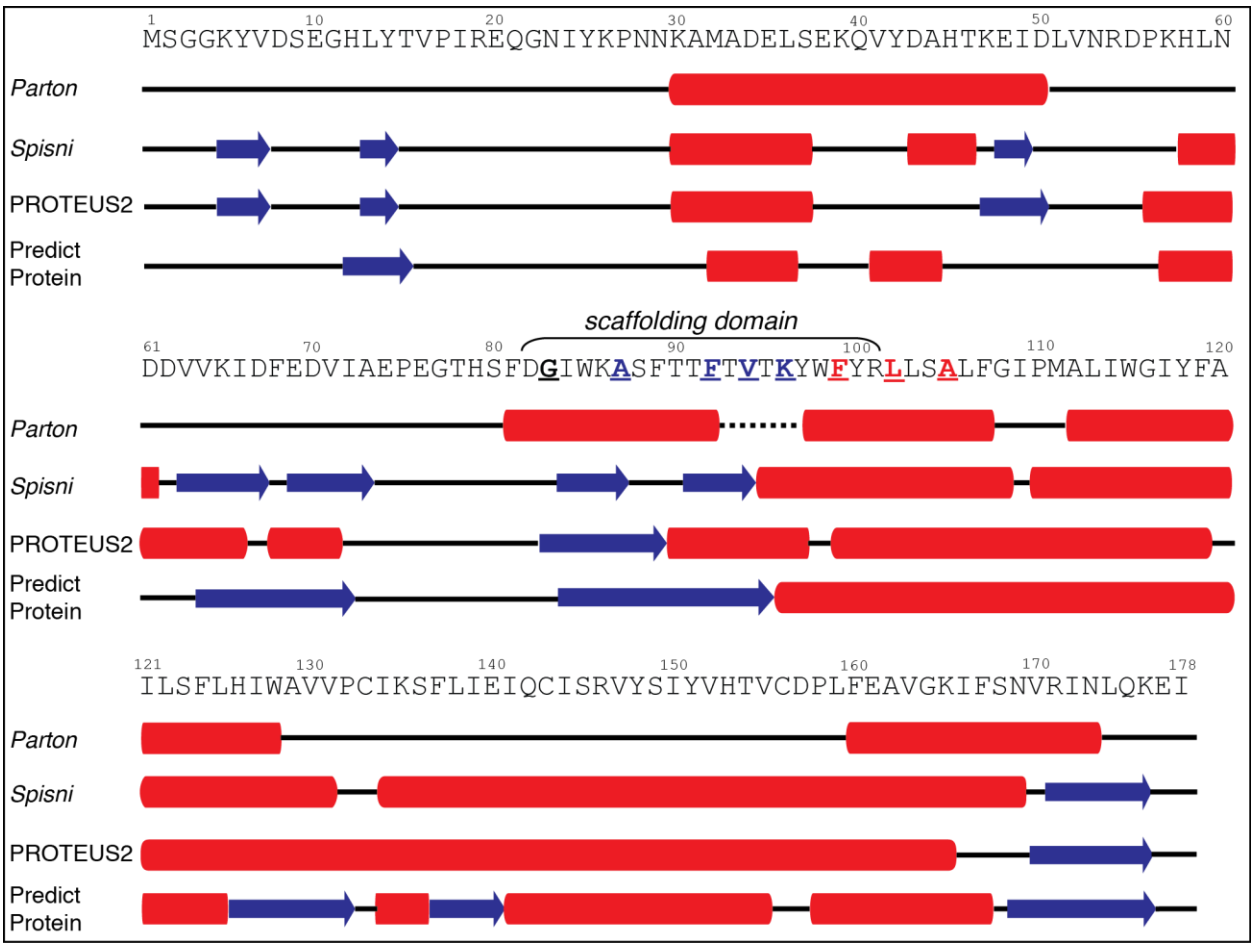

Figure 4.10 Schematic illustrations of secondary structure predictions for full-length Cav1.

*Parton et al. (top row) have presented an all-helical model based on primary sequence analysis [27]. In a bioinformatics study, Spisni et al. used an array of algorithms to obtain a model in which residues 84-94 form an anti-parallel β-hairpin [215]. In our hands, the PSIPRED (**Figure 4.9**), PROTEUS2, and PredictProtein algorithms predict significant β-sheet secondary structure, although with slightly varying distributions. Predictions of β-structure within the CSD by the PSIPRED, PredictProtein and Spisni models quite closely match our experimental observations on the CSD-containing Cav1 fragments, as marked by color-coding on the primary sequence (also see **Figure 4.8 and Figure 4.9**).*

4.4 DISCUSSION

Based on a combination of ssNMR, CD, and FTIR measurements, we have characterized the secondary structure of Cav1 fragments, while bound to cholesterol-rich model membranes. These experiments were designed to examine the CSD, as it harbors many Cav1 functionalities. The $Cav_{82\text{-}109}$ peptide, incorporating both the CSD and the first half of the IMD, was found to contain a mixture of β-sheet and α-helical structure. The latter was localized within the putative first α-helix of the helical hairpin of the IMD, consistent with previous work and structural models and predictions [208-211]. More surprisingly, our observations indicate a β-strand conformation for the majority of labeled residues within the CSD, based on ssNMR shift analysis and inter-residue contacts, as well as the detection of mixed secondary structure by CD and FTIR. Computational analysis of the primary sequence predicts that part of the CSD prefers a β-sheet conformation, and matches both the location of the experimental transition to α-helix and the observed lack of helicity in short peptides [208, 215]. **Figure 4.11** contains a graphical schematic that summarizes these observations. The observed helicity of the first half of the IMD is consistent with reports proposing models where the P110 forms the turn in a 're-entrant' helix stabilized by interacting 'smaller' residues in both its helical segments [27, 200, 201]. This domain's 'wedge' shape that fails to traverse the entire bilayer may well contribute to the remarkable membrane curvature of the caveolae [242].

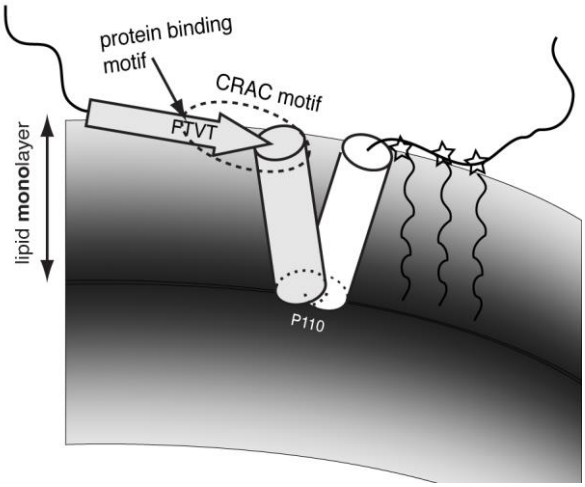

Figure 4.11 Schematic illustration of secondary structure distribution of membrane-bound Cav1.

The wedge-shaped α-helical hairpin is too short to traverse the entire bilayer, and occupies a single monolayer. The CSD is shown as largely β-stranded, with the putative CRAC motif present at the transition between β- and α-structure. The C-terminal palmitoylation sites are also indicated. Note that the exact orientation and structure of the β-strands remain uncertain, and may well reflect an antiparallel β-hairpin instead of a single strand as shown in the figure.

4.4.1 Existing Structural Data and Models for the CSD

Previous reports on Cav and the CSD fail to reach a consensus on the structure of the CSD. Studies by Fernandez et al via CD and sequence analysis of a $Cav_{1\text{-}101}$ fragment in solution proposed the presence of an α-helix spanning residues 79-96 [210]. Solution NMR and CD on solubilized $Cav_{82\text{-}101}$ also indicated a predominantly helical structure spanning residues Ile84-Tyr97 [212]. However, the latter experiments were largely done in presence of the helicity-enhancing co-solvent TFE, which was reportedly required to avoid aggregation. CD spectroscopy on $Cav_{83\text{-}102}$ peptides associated with POPC vesicles [32] and $Cav_{82\text{-}101}$ peptides in

DPC micelles [208] indicated only a partially α-helical structure for the isolated CSD. CD on a longer construct Cav_{82-109} (also employed here) revealed a large increase in α-helicity and solution NMR on the micelle-associated peptides localized the α-helical structure to residues 83-88 and 93-97 of the CSD and 102-108 within the IMD [208]. In some of these studies computational analysis of the primary sequence supported a predominantly α-helical CSD conformation, even leading to a predominantly helical full-length model of Cav [27, 210]. On the other hand, an in silico study by Spisni et al has proposed an anti-parallel β-hairpin structure in residues 84-94 of the CSD, which is remarkably consistent with our findings [215].

Despite the lack of consensus, our observations on membrane-bound peptides recapitulate a number of features reported above. We observe that the peptides are characterized by a partially α-helical conformation that is most pronounced in the C-terminal residues (beyond the CSD proper). The key difference in this study is the experimental identification of a β-stranded conformation within the CSD (in cholesterol-rich bilayers), and the fact that we site-specifically delineate the location of the different secondary structures. Our ssNMR results show that the observed mixed secondary structure content, with substantial non-helical structure, is not simply due to sample heterogeneity (with some peptide largely helical but others unstructured), but rather reflects the formation of distinct domains.

Our work is also distinct in that we probe the structure of the peptide in the presence of high levels of cholesterol, to mimic the caveolar membrane and to facilitate examination of the CSD-cholesterol interaction. MAS and static ssNMR experiments clearly show that the lipid mixture responds to the integration of the peptides without generating peptide-induced heterogeneity in the lipids. Thus, we see a membrane-bound peptide and not simply the co-precipitation of non-interacting peptides and lipids.

4.4.2 Scaffolding Domain Functionalities

In the following sections, we will examine our observations in view of the different roles ascribed to the CSD, in an attempt to place our results in context and discuss potential implications for each of these CSD functionalities.

1. CSD binding partners. One critical caveolin function localized within the CSD relates to interactions with its numerous binding partners, such as protein G α-domains and Src family kinases [206]. Within the partner protein, this is mediated by the CBM, with a consensus sequence (ΦXΦXXXXΦXXΦ) rich in aromatic residues (Φ). The CBM of the protein G α-subunit forms a β-hairpin (**Figure 4.12a**) [206]. Within the CSD, residues F92TVT95 are involved, which adopt a β-stranded conformation in our samples. In various studies [33, 206, 243];Li, 1996, p09577;Toya, 1998, p09575}, isolated short CSD peptides appear to emulate the effect of Cav binding to its binding partners, suggesting that the Cav tertiary structure is not required. Often a membrane is not present in those studies, raising the question of whether a membrane-bound CSD conformation is at all involved and thus how our data correlate to this CSD role.

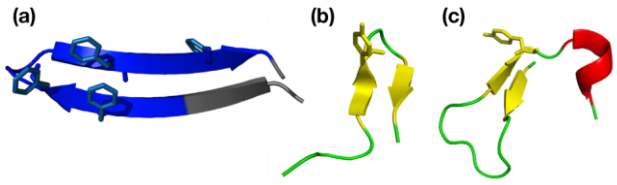

Figure 4.12 Secondary structure motifs in proposed homologous protein domains.

(a) Structure of the aromatic-residue-rich caveolin binding sequence as originally identified in Gα proteins [206], featuring a β-hairpin structure. (b,c) β-strand rich structures found among four putative CRAC motifs [244] in the PDC-109 crystal structure [245]. See ***Figure 4.13*** *for the complete protein context and additional details.*

2. Oligomerization. The CSD is also involved in the formation of Cav homo-oligomers, and oligomerization is a critical functionality as it appears to be required in the formation of caveolae. We note that a membrane-embedded β-stranded CSD may be prone to self-assemble into β-sheets in order to eliminate membrane-exposed non-hydrogen-bonded backbone sites. Indeed, a bioinformatics study by Spisni et al has suggested that such a β-hairpin structure in residues 84-94 could play a role in homo-oligomer interactions [215]. It has also been noted that Cav1 appears to have a motif in residues 92-100, coinciding with the CSD, which strongly resembles the CBM [206]. This may indicate a mechanism for Cav self-oligomerization that is analogous to its recognition of the binding partner proteins. If so, this seems consistent with a potential CSD β-hairpin structure, given the hairpin conformation of the CBM (see above).

Our current data do not inform us on the oligomeric (or monomeric) nature of the peptides under investigation. Long-mixing ^{13}C-^{13}C experiments could be employed to identify intermolecular interactions and detect and characterize the supramolecular assemblies [246, 247]. However, the long-mixing experiments on the current samples have failed to produce data consistent with intermolecular interactions, including experiments performed on samples prepared from mixtures of peptides p1 and p2 (not shown). This may be due to the sparse labeling and the relatively low signal intensities for some of the samples. More indirectly, the apparent rigidity of the peptides and pronounced effect on membrane fluidity could be indicative of the formation of peptide oligomers. However, this may also be explained in part by the high level of cholesterol present. Note that our data do show clear effects on the membrane dynamics and that the peptides are homogeneously associated with the membrane, resulting in a single population of lipid NMR as well as peptide signals. This indicates that any oligomerization that

has occurred is taking place in a membrane-bound state, and thus may be mimicking the homo-oligomer-formation that Cav displays in cholesterol-rich caveolae.

3. Cholesterol binding. The binding of cholesterol is thought to be mediated by residues V_{94}TKYWFYR_{101} that form a putative CRAC motif. The original CRAC motif was identified in the translocator protein TSPO [248]. Subsequent solution NMR studies of various CRAC motifs have proposed a predominantly α-helical CRAC structure [208, 211, 213], although other work suggested a lower level of α-helicity [32]. This difference has been credited to a loss of helical structure due to truncation in the shorter peptides used in the latter studies [208]. The above structural work involved short peptides lacking protein context, but a recent study reported on CRAC motifs in the protein PDC-109, for which a known homologous structure exists [244]. While this seemingly provides much-needed high-resolution structures of these motifs in a native context, a more detailed inspection reveals an array of different structures (**Figure 4.12b and c and Figure 4.13**). This may act as a reminder that these are *putative* motifs and that some may not be functional cholesterol binding domains. Nonetheless, while PDC-109 fails to provide an unambiguous structure for the CRAC motif, it may provide examples of possible β-rich CRAC conformations (**Figure 4.12b and c**). Both our experimental and computational data indicate a specific transition of β-sheet to α-helical structure to occur within the CRAC motif, near the motif's conserved central Tyr97, reminiscent of earlier computational predictions [215]. Intriguingly, at least one of the PDC-109 CRAC motifs (**Figure 4.12c**) shows a similar motif, with the central Tyr at the C-terminal end of a β-hairpin and near the start of a flanking helical segment. The observed sheet-helix transition is also reminiscent of the tendency for CRAC motifs to occur directly adjacent to transmembrane helices [249].

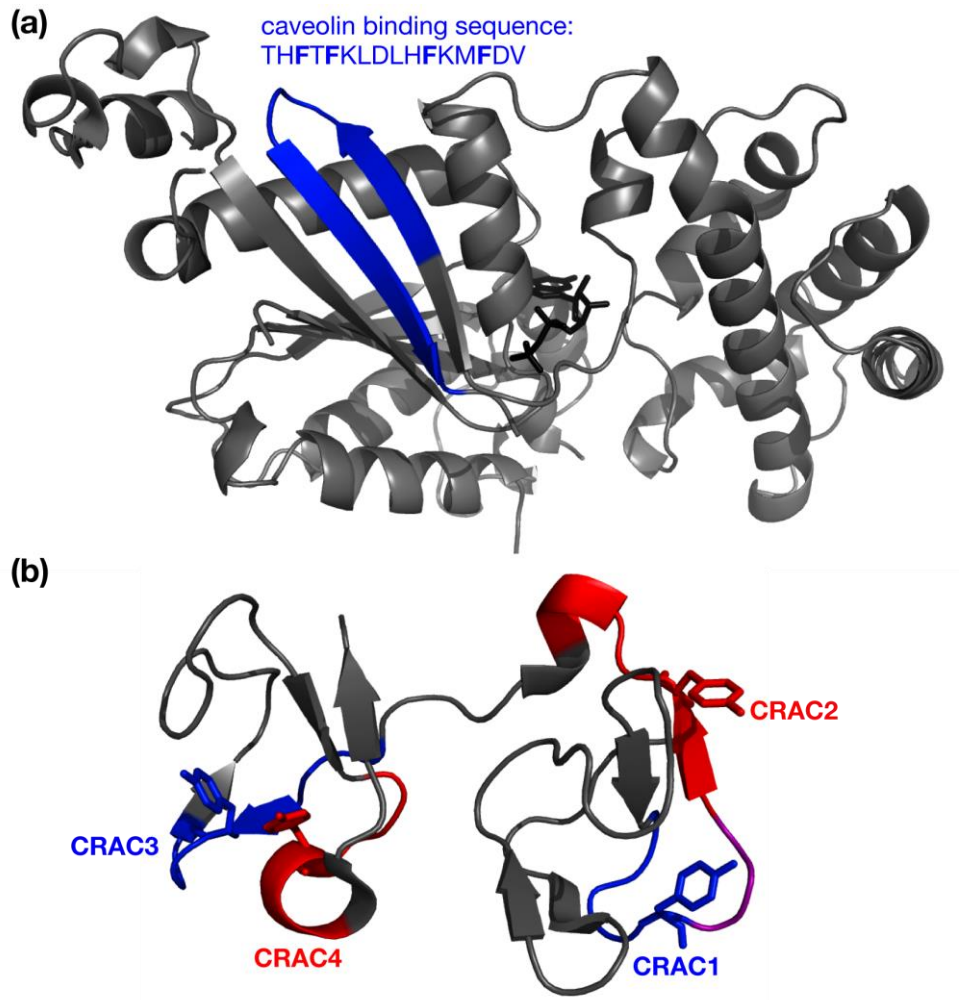

Figure 4.13. Structural variations in caveolin binding and cholesterol recognition motifs in previously solved structures.

(a) Structure of rat Giα1 with GDP bound (black sticks) [250], a binding partner of caveolin [251]. The caveolin binding sequence is shown in blue, and has apparent homology to a motif within caveolin itself [206]. (b) Structure of PDC-109, showing its four putative CRAC motifs. The CRAC motifs are alternatingly colored blue and red (overlapping residues in purple). The Tyr residues at the heart of each CRAC motif are shown in stick representation. Note the variation in (secondary) structure between the different motifs, but also that the structure is not necessarily α-helical. Two motifs contain anti-parallel β-sheet structures, one of which (CRAC2) is flanked by a short helix as commonly found for CRAC motifs (although in this case not transmembrane). These figures were prepared using PyMol (Schrödinger).

Given the reported strong interactions between caveolin and cholesterol, one key point of interest will be to characterize this interaction in more detail in future work. Interestingly, in the current work we have not seen large effects for the peptide NMR signals of the omission of cholesterol (**Figure 4.3**). This suggests that the observed secondary structures are not directly correlated to cholesterol-binding, raising the question of whether this indicates a more subtle effect of cholesterol binding, the need for other parts of Cav, or a role for other lipids enriched in caveolae (e.g. sphingomyelin).

The above shows that there have been indications of a role for β-stranded conformations in CSD, and that such a structure may make sense from a functional point of view. One seemingly common feature is the presence of β-hairpins, as also suggested by Spisni et al, who proposed that an anti-parallel β-hairpin is formed by residues 84-94 [215]. Our ssNMR data reveal the localized β-conformation, but having thus far failed to observe long-distance inter- or intramolecular interactions, it remains uncertain whether this represents a single β-strand or shorter β-strands within e.g. a β-hairpin. The relatively sparse labeling applied here makes it impossible to exclude the hairpin motif based on these data, as it is quite possible that labeled sites in a β-hairpin would be too far apart. Whether in a hairpin or not, it may well be that the β-strands play a role in the oligomerization process, as membrane-embedded strands could be driven to self-assemble to prevent exposure of the peptide backbones to the hydrophobic membrane environment [215].

4.5 CONCLUSION

Our experimental and computational analysis of the structural features of Cav1 fragments seems to indicate a more significant propensity for β-sheet structures with the CSD than apparent from the existing literature. We have discussed that such β-rich structures could be consistent with its various functionalities, while highlighting the need for further characterization. One important consideration is that the caveolins fulfill a variety of functions, which could be independent of each other, occur in different locations within the cell, and may involve different Cav conformations. For instance, it has been pointed out that palmitoylation is non-essential for the localization in caveolae, but that it is critical for the cholesterol trafficking by Cav [204]. It seems likely that the Cav conformation within caveolae could be significantly different from its structure in a cholesterol-bound cytosolic state. Thus, it is important to investigate Cav structure in different contexts, including cholesterol-rich membranes as examined here. Thus far, we lack a clear understanding of the structural effects of cholesterol binding, as we did not observe a large structural change in the presence of cholesterol. Clearly, additional experimental and structural data will be critical to elucidate the conformation and oligomeric states of Cav, and to provide an understanding of the effect of disease-causing mutations and the structural underpinning of its many functional roles within the cell.

4.6 ACKNOWLEDGEMENTS

We thank Mike Delk for technical assistance and Ron Wetzel for the use of laboratory equipment.

5.0 THE BIG PICTURE SUMMARY

Amyloidogenic proteins cause more than 20 human diseases [1]. At least ten human disorders are caused by polyglutamine expansions [57]. For many of these disorders the toxic species is unknown. However, it is clear that protein aggregation is associated with disease progression. Understanding the mechanism of amyloid formation will aid in discovery of inhibitors and new therapeutics against these diseases. Throughout the aggregation mechanism of amyloidogenic proteins, structural changes are taking place: from protein misfolding, to formation of oligomers, to assembly of rigid, compact, β-rich amyloid fibrils. I have shown that through use of MAS ssNMR, we can use model htt N-terminal fragment peptides to delve into the structure of the polyQ domain and its flanking segments in the end stage fibrils. In **Chapter 2.0** , we identified the boundary of the polyQ amyloid core, which extends up to the PPII-structured oligoPro domain. The polyQ flanking segments have opposing impacts on the aggregation kinetics. In solution, the oligoPro domain has been shown to induce PPII structure into the polyQ domain, suppressing nucleation, and thus aggregation. However, once in mature fibrils, the PPII oligoPro causes minimal disruption of the polyQ amyloid core and the structural transition from β-sheet to PPII structure is remarkably abrupt. The junction has a higher mobility than the rigid amyloid core and is more solvent exposed. This site actually serves as a binding site for certain htt-binding proteins, some of which can bind to and destabilize fibrils [116, 118, 149]. This may suggest that this site could serve as a potential binding site for drug compounds.

The α-helix of the htt^{NT} participates in htt^{NT}-htt^{NT} intermolecular interactions to accelerate fibril formation. In **Section 1.3.1**, I showed that α-helicity remains in the htt^{NT} in the mature fibril aggregates. The α-helix is more mobile and solvent-exposed than the amyloid core. In **Chapter 2.0** , I showed that this α-helix displays solvent-coupled dynamics. While the htt^{NT} retains a structured α-helix, it is dynamic and is seemingly packed relatively loosely, which is consistent with a molten-globule-like assembly observed in other oligomeric amyloid intermediates [153-156]. Incorporating different isotopic labeling schemes to full-length htt exon 1 could give insight regarding the htt^{NT}-htt^{NT} interactions and the proximity of the α-helices in mature fibrils. With new insights into the structure and dynamics of the flanking segments that influence fibril formation, we may better understand the aggregation mechanism and also binding characteristics and effects of htt-binding proteins.

The architecture of the polyQ core, itself, in mature fibrils is integral to understanding its assembly. In all ssNMR studies of polyQ-containing fibrils, unique chemical shift patterns have been observed for amyloid core Gln residues. In **Section 1.3.2**, I have shown these chemical shifts to be unique based on an analysis of all Gln chemical shifts reported in the BMRB and characteristic of polyQ core Gln residues [25]. These shifts are indicative of two major Gln conformers with distinct geometries. In **Chapter 3.0** , I have shown that we can obtain structural constraints and relaxation parameters from MAS ssNMR techniques, which have given us new insight into the origin of this "signature motif." I have proposed a new structural model of the arrangement and side chain conformations of the Gln conformers in the amyloid core. The configuration of turns between β-strands is also part of the fibril architecture and assembly. In **Chapter 3.0** , I discussed the usage of β–turn enhancing Pro-Gly motifs in the middle of a polyQ stretch. Torsion angle constraints and chemical shift analysis by MAS ssNMR have shown how

polyQ is able to accommodate β-II turns into mature fibrils and provides additional support for a polyQ aggregation mechanism involving β-hairpin structures. This work suggests feasibility of studies of polyQ in atomic detail by MAS ssNMR. Further refinement of structural features of Gln residues in the β-strands and turns within the fibril will give a clearer model of the polyQ fibril assembly.

Cav1 forms homo-oligomers and associates with cholesterol-rich membranes through its CSD to form caveolae. In **Chapter 4.0** , I discussed the observation of β-sheet structure in the CSD of Cav1 fragment peptides based on multiple spectroscopic and computational techniques. This contrasts with previous studies, which did not use a cholesterol-rich lipid environment that are likely to impact the conformation. Further studies of intermolecular protein-protein and protein-cholesterol interactions should be pursued to probe the oligomeric structures of Cav1 and how it interacts with cholesterol in the membrane. Additionally, the impact of the protein context should be explored, perhaps by studying structural features in the full-length Cav1 protein.

APPENDIX A

SUPPORTING TABLES FOR CHAPTER 2

Table A.1 Detailed experimental conditions of NMR experiments shown in Chapter 2.0

Abbreviations: NS, number of scans per t_1 point; Temp., temperature; MAS, magic angle spinning rate; RD, recycle delay; TPPM, ^{1}H decoupling power during evolution and acquisition (using two-pulse phase modulation scheme).

1D Spectra									
Fig.	**Sample (refer to Table 2.1)**	**Experiment**	**NS**	**Temp (K)**	**MAS (kHz)**	**RD (s)**	**TPPM during acq. (kHz)**	**T_2 filter time (ms)**	**^{1}H-^{1}H Mixing (ms)**
2.2a, 2.3a,d, 2.4	LQP-labeled	^{1}H-^{13}C CP	1024	275	9.8	2.8	83	NA	NA
2.2b	htt^{NT}Q$_{30}$P$_{10}$K$_2$ *	^{1}H-^{13}C CP	2662	275	10	3	83	NA	NA
2.3b,c	LQP-labeled	^{1}H T_2 filter	3072	275	9.8	2.8	83	3.0	0
2.3c	LQP-labeled	^{1}H T_2 filter	3072	275	9.8	2.8	83	3.0	1
2.3c	LQP-labeled	^{1}H T_2 filter	3072	275	9.8	2.8	83	3.0	2
2.3c	LQP-labeled	^{1}H T_2 filter	3072	275	9.8	2.8	83	3.0	3
2.3c	LQP-labeled	^{1}H T_2 filter	3072	275	9.8	2.8	83	3.0	4
2.3b,c	LQP-labeled	^{1}H T_2 filter	3072	275	9.8	2.8	83	3.0	7
2.3d	LQP-labeled	^{1}H-^{13}C DP	1024	275	9.8	2.8	83	NA	NA
2.8a	MA-labeled	^{15}N T_1	1024	315	22	3	83	NA	NA
2.8a	MA-labeled	^{15}N T_1	1024	273	19	3	83	NA	NA
2.8b	[U-^{13}C,^{15}N-Q10]-K$_2$Q$_{11}$PGQ$_{11}$D$_2$	^{15}N T_1	790	275	22	6	83	NA	NA
2.8c,d	MA-labeled	N-H dipolar coupling	2048	287	10	3	83	NA	NA
2.8c,d	MA-labeled	N-H dipolar coupling	2048	250	10	3	83	NA	NA
2.8e	[U-^{13}C,^{15}N-Q10]-K$_2$Q$_{11}$PGQ$_{11}$D$_2$	N-H dipolar coupling	2048	275	10	3	83	NA	NA
2.9a	1:1 MF-labeled/LAQ-labeled	^{1}H-^{13}C CP	256	287	9.8	3	83	NA	NA

2.9a	1:1 MF-labeled/LAQ-labeled	^{1}H-^{13}C DP	256	287	9.8	3	83	NA	NA
2.9b	1:1 MF-labeled/LAQ-labeled	^{1}H-^{13}C CP	256	265	9.8	3	83	NA	NA
2.9b	1:1 MF-labeled/LAQ-labeled	^{1}H-^{13}C DP	256	265	9.8	3	83	NA	NA
2.9c	1:1 MA-labeled/LQP-labeled	^{1}H-^{13}C CP	256	287	9.8	3	83	NA	NA
2.9c	1:1 MA-labeled/LQP-labeled	^{1}H-^{13}C DP	256	287	9.8	3	83	NA	NA
2.9d	1:1 MA-labeled/LQP-labeled	^{1}H-^{13}C CP	256	270	9.8	3	83	NA	NA
2.9d	1:1 MA-labeled/LQP-labeled	^{1}H-^{13}C DP	256	270	9.8	3	83	NA	NA
2.4b	LKSQ-labeled	^{1}H-^{13}C CP	4096	275	13	3.5	83	NA	NA
2.4b	LKSQ-labeled	^{1}H-^{13}C DP	4096	275	13	3.5	83	NA	NA

2D Spectra

Fig.	Sample (refer to Table 1 in main text)	Experiment	NS	Temp (K)	MAS (kHz)	RD (s)	TPPM during acq. (kHz)	t_1 evol. (μs)	Mixing (ms)
2.1a	LQP-labeled	DARR 2D	64	275	9.8	2.8	83	422x33.11	8
2.1b	LQP-labeled	DARR 2D	72	275	9.8	2.8	83	370x36.78	15
2.1c	LAQ-labeled	DARR 2D	128	276	10	2.8	83	240x36.78	8
2.1d	LKSQ- labeled	DARR 2D	96	275	13	3	83	832x19.23	25
2.5a,b, 2.6	U-^{13}C,^{15}N-htt exon 1	DARR 2D	256	275	10	2.6	83	448x35.60	15
2.5b, 2.6	LQP-labeled	DARR 2D	72	275	9.8	2.8	83	370x36.78	15
2.5b, 2.6	LKSQ- labeled	DARR 2D	96	275	13	3	83	832x19.23	25
2.7a	MA-labeled	DARR 2D	64	275	10	2.8	83	422x33.1	8

* Natural abundance signals from $htt^{NT}Q_{30}P_{10}K_2$ aggregates that lacked isotopic labeling.

Table A.2 ^{13}C and ^{15}N chemical shift assignments of residues isotopically labeled in $htt^{NT}Q_{30}P_{10}K_2$ peptide fibrils.

The uncertainty in the chemical shifts is ± 0.1-0.3 ppm unless otherwise stated. ^{13}C referencing is relative to aqueous DSS (see Experimental Procedures section). These data are also available online at the Biological Magnetic Resonance Data Bank (BMRB), via BMRB accession number 25146.

Res.[a)]	CO	Cα	Cβ	Cγ	Cδ(1)	Cδ2	Cε	N	Nε2
A2	178.2	52.7	19.1						
L4	178.4	58.0	41.4	27.0	25.3	24.2			
K6	179.9	59.5	32.6	25.8	29.6		42.1		
L7	178.0	57.9	42.0	26.9	25.2	24.0		121.6	
M8	178.9	58.0	32.1	32.3			17.0	118.7	
A10	180.3	55.0	18.0					123.2	
F11	-	61.2	39.3	131.4					
L14	177.0 ± 0.4	55.7 ± 0.4	42.1	26.6	26.6	23.3			
S16a	173.0 ± 0.5	56.8	65.2						
S16b	173.0 ± 0.5	58.9	62.8						
F17a	175.8	57.1	-	131.7					
F17b	174.3	56.6	-	131.7					
Q18a	175.6	55.7	34.2	34.4	178.8				
Q18b	174.2	54.7	30.9	30.7	177.8				
Q19a	176.0	56 .0	34.2	34.2	178.6				
Q19b	174.2	53.9	31.1 ± 0.4	30.6	177.6				
Q46a	175.1	56.0	34.3	34.0	178.7				
Q46b	174.2	54.1	31.6	30.2	178.3				
Q46c	-	-	-	34.1	179.8				
Q47c1	172.5	53.7	29.1	33.4	180.4			123.1	111.4
Q47c2	173.0 ± 0.5	53.0	30.3	34.3	178.6			117.7	107.7
P48	174.2	61.3	30.5	27.3	50.5			136 ± 1	
Pro NA	-	61.2	30.6	26.8	50.3				

APPENDIX B

SUPPORTING TABLES FOR CHAPTER 3

Table A.3 Detailed experimental conditions of 1D and 2D NMR experiments shown in Chapter 3.0

Abbreviations: NS, number of scans per t_1 point; Temp., temperature of cooling gas; MAS, magic angle spinning rate; RD, recycle delay; TPPM, ^1H decoupling power during evolution and acquisition (using two-pulse phase modulation scheme).

1D Spectra							
Figure	**Sample**	**Experiment**	**NS**	**Temp (K)**	**MAS (kHz)**	**RD (s)**	**TPPM during acq. (kHz)**
3.5a	[U-^{13}C,^{15}N-Q10]-$K_2Q_{11}PGQ_{11}D_2$	^{15}N $R_{1\rho}$	10,240	296	60	2	15
3.5b	[U-^{13}C,^{15}N-Q10]-$K_2Q_{11}PGQ_{11}D_2$	^{15}N R_1	790	275	22	6	83
3.5c	[U-^{13}C,^{15}N-Q10]-$K_2Q_{11}PGQ_{11}D_2$	^{13}C $R_{1\rho}$	2048	296	60	2	15
3.5d	[U-^{13}C,^{15}N-Q10]-$K_2Q_{11}PGQ_{11}D_2$	^{13}C R_1	2048	296	60	2	15
3.6a,b	pG1	HCCH	1536	276	6	3	83
3.6c,d	Q15	HCCH	2048	275	6	2.7	83
3.7c	pG2	CONCX	2048	276	10	3	83
3.9	PG1	^{15}N R_1	1024	275	22	6	83

2D Spectra									
Figure	**Sample**	**Experiment**	**NS**	**Temp (K)**	**MAS (kHz)**	**RD (s)**	**TPPM during acq. (kHz)**	**t_1 evol. (μs)**	**Mixing (ms)**
3.1a	[U-^{13}C,^{15}N-Q6]-$K_2Q_{30}K_2$	DARR 2D	96	275	10	3	83	550x22.07	25
3.1b	Q15	DARR 2D	128	275	10	2.8	83	512x38.95	8
3.2b	pG1	DARR 2D	64	276	10	2.8	83	450x36.78	8
3.2c,d	pG1	NCO	256	276	10	2.8	83	98x200.0	NA

3.2e	htt exon 1	NCO							
3.2f	htt exon 1	NCA							
3.4a,c	pG1	NCCN	256	276	10	3	83	100x40.00	0.8
3.7a	pG2	NCA	536	276	10	3	83	50x200.0	NA
3.7b	pG2	NCOCX	512	276	10	3	83	50x200.0	15
3.7d,e, 3.8a	pG2	DARR 2D	96	276	10	2.8	83	400x36.78	8
3.8b	PG1	DARR 2D	344	275	10	3	83	334x30.00	10
3.10a,b 3.11a	PG1	PDSD 2D	104	275	8	2.5	83	332x30.00	500
3.10b, 3.11b	PG1	PDSD 2D	104	275	8	2.5	83	332x30.00	250
3.10b, 3.11c	PG1	DARR 2D	104	275	8	2.5	83	332x30.00	100
3.13	PG1	NCCN	352	275	10	2.5	83	100x40.00	0.8

Table A.4 ^{13}C and ^{15}N chemical shift assignments of residues isotopically labeled in $K_2Q_{30}K_2$, $D_2Q_{15}K_2$, $K_2Q_{11}pGQ_{11}K_2$, and $K_2Q_{11}PGQ_{11}D_2$ peptides.

Chemical shifts are accurate to +/- 0.1-0.3 ppm unless otherwise stated.

Res.	C'	Cα	Cβ	Cγ	Cδ	N	Nε2
$K_2Q_{30}K_2$ fibrils [21]							
Q6a	175.9	56.1	34.2	34.2	178.6	--	--
Q6b	173.8	54	31.7	30.2	177.4	--	--
Q6c	--	--	--	34.2	180	--	--
$D_2Q_{15}K_2$ fibrils							
Q6a	175.9	56.1	34.2	34.1	178.8	--	--
Q6b	173.9	53.9	31.7	30	177.5	--	--
Q6c	--	--	--	33.8	180.4	--	--
$K_2Q_{11}pGQ_{11}K_2$ fibrils							
Q8a	176.1	56.3	34.2	34.2	178.8	119.1	108.5
Q8b	173.9	53.9	31.7	30	177.5	116.9	105.9
Q8c	--	--	--	33.7	180.3	--	111.4
Q9a	176.1	56.3	34.2	34.2	178.8	119.1	108.5
Q9b	173.9	53.9	31.7	30	177.5	117.1	105.9
Q9c	--	--	--	33.7	180.3	--	111.4
Q13	173.9	53.9	31.9	30.4	178.1	119	107.1
G15a	174	46.1					108.2
G15b	173.4	45.7					114.2
Q16	175.7	55.7	34.2	34.2	178.7	120.3	109.9
Q13/16	--	54.9	31.5	34.2	180.4	121.1	111.6 +/- 0.7
$K_2Q_{11}PGQ_{11}D_2$ fibrils							
Q10a[25]	175.8	55.8	34	33.9	178.5	--	--
Q10b[25]	173.8	53.6	31.5	29.9	177.5	--	--
Q10c[25]	--	--	--	33.2	180	--	--
Q13a	173.1	53.5	29.8 +/- 0.5	32.5	177.6	119	104.2
Q13b	172.8	53.7	30.3	33.6	180.1	119	109.1
P14a	177.4	66.2	31.2	28.3	50.2	135.6	
P14b	176.6	64	31.9	27.8	50.1	133.5	
P14c	175.1	63.2	34.1	24.9	50.1	131.8	
G15a	174	46				111	
G15b	174	45.9				112.2	
G15c	174	46.1				107	

APPENDIX C

SUPPORTING TABLES FOR CHAPTER 4

Table A.5 Detailed experimental conditions of 1D and 2D NMR experiments shown in Chapter 4.0

Abbreviations: NS, number of scans per t_1 point; Temp., temperature of cooling gas; MAS, magic angle spinning rate; RD, recycle delay; TPPM, ^1H decoupling power during evolution and acquisition (using two-pulse phase modulation scheme).

1D Spectra							
Figure	**Sample**	**Experiment**	**NS**	**Temp (K)**	**MAS (kHz)**	**RD (s)**	**TPPM during acq. (kHz)**
4.2c	1:1 POPC (6.3mg) /Chol (3.2mg)	^{1}H-^{13}C CP	2560	283	8	3.5	83
4.2a	10mol% na-Cav$_{82-109}$ (6mg) in 1:1 POPC (6.2mg)/Chol (3.1mg)	^{1}H-^{13}C CP	2560	283	8	3.5	83
4.2b	10mol% na-Cav$_{83-102}$ (6.75mg) in 1:1 POPC (9.14mg)/Chol (4.59mg)	^{1}H-^{13}C CP	5120	300	8	3.5	50
4.2d, 4.3d	POPC (5mg)	^{31}P 1D	3072	300	0	3.5	50
4.2e	10mol% p1 in POPC (4.47mg)	^{31}P 1D	3072	300	0	3.5	50
4.2f	1:1 POPC (6.3mg)/Chol	^{31}P 1D	2048	300	0	3.5	50
4.2g	10mol% na-Cav$_{83-102}$ in 1:1 POPC (9.14mg)/Chol	^{31}P 1D	6144	300	0	3.5	50

4.2h, 4.3f	10mol% p1 in 1:1 POPC (7.08mg)/Chol	^{31}P 1D	2048	300	0	3.5	50
4.3a	10mol% p1 (6.8mg) in 1:1 POPC (7.08mg) /Chol (3.56mg)	^{1}H-^{13}C CP	1024	283	8	3	71.5
4.3b	2mol% p1 (2.15mg) in 1:1 POPC (12.16mg)/Chol (6.11mg)	^{1}H-^{13}C CP	4096	283	8	3.5	83
4.3c	10mol% p1 (2.15mg) in POPC (4.47mg)	^{1}H-^{13}C CP	4096	283	5	3.5	83
4.3e	2mol% p1 in 1:1 POPC (12.16mg) /Chol	^{31}P 1D	2048	300	0	3	50

2D Spectra

Figure	**Sample**[a)]	**Experiment**	**NS**	**Temp (K)**	**MAS (kHz)**	**RD (s)**	**TPPM during acq. (kHz)**	**t_1 evol. (μs)**	**Mixing (ms)**
4.6, 4.3	p1	DARR 2D	152	283	8	3.5	83	256x31.25	25
4.6b	p2	DARR 2D	152	283	8	3.5	83	256x31.25	25
4.7a	p1	PDSD 2D	104	283	8	3	71	256x31.25	400
4.7b	p2	PDSD 2D	152	283	8	3	83	256x25.00	400
4.6a	p3	DARR 2D	144	283	8	3	83	283x32.00	25

[a)] All 2D spectra were acquired on 10 mol% of the respective peptide in 1:1 (molar ratio) POPC/cholesterol membranes.

Table A.6 ^{13}C chemical shift resonances for Cav1 fragment peptides obtained by solid state MAS NMR.

*Entries for residues A87 and V94 include the shifts of a minor secondary population (ref. **Chapter 3.6**). Chemical shifts are referenced relative to aqueous DSS through indirect referencing to external adamantane. The last column summarizes the secondary structure assignments (see **Chapter 4.0**).*

	Res.	Cα	Cβ	Cγ	Cδ(1)	Cδ2	Cε	CO	
Caveolin Scaffolding Domain	G83	46.47±0.04						170.50±0.07	-
	A87	50.6±0.2	23.6±0.2					176.0±0.2	β
		50.61±0.01	22.60±0.06					176.0±0.2	β
	F92	56.3±0.2	43.9±0.2	139.2±0.1	132.40±0.01		131.50 ± 0.03	174.3±0.2	β
	V94	60.5±0.2	35.5±0.2	22.0±0.1				173.9±0.3	β
			32.0±0.4	21.8±0.1					
	K96	55.0±0.2	37.1±0.3	25.9±0.5	29.9±0.2		42.5 ± 0.1	174.1±0.3	β
	F99		38.8±0.1	131.20±0.09					α
IMD	L102	58.3±0.1	41.6±0.2	27.30±0.09	25.10±0.09	23.6±0.2		178.1±0.2	α
	A105	54.5±0.2	18.90±0.09					179.2±0.4	α

BIBLIOGRAPHY

1. Chiti, F. and C.M. Dobson, *Protein misfolding, functional amyloid, and human disease.* Annu Rev Biochem, 2006. **75**: p. 333-66.
2. Nagai, Y., et al., *A toxic monomeric conformer of the polyglutamine protein.* Nat Struct Mol Biol, 2007. **14**(4): p. 332-40.
3. Bodner, R.A., D.E. Housman, and A.G. Kazantsev, *New directions for neurodegenerative disease therapy: using chemical compounds to boost the formation of mutant protein inclusions.* Cell Cycle, 2006. **5**(14): p. 1477-80.
4. Sunde, M., et al., *Common core structure of amyloid fibrils by synchrotron X-ray diffraction.* J Mol Biol, 1997. **273**(3): p. 729-39.
5. Li, J., et al., *Amyloid-like Fibrils from a Domain-swapping Protein Feature a Parallel, in-Register Conformation without Native-like Interactions.* Journal of Biological Chemistry, 2011. **286**(33): p. 28988-28995.
6. Luhrs, T., et al., *3D structure of Alzheimer's amyloid-beta(1-42) fibrils.* Proc Natl Acad Sci U S A, 2005. **102**(48): p. 17342-7.
7. Wasmer, C., et al., *Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core.* Science, 2008. **319**(5869): p. 1523-6.
8. Fandrich, M., M. Schmidt, and N. Grigorieff, *Recent progress in understanding Alzheimer's beta-amyloid structures.* Trends Biochem Sci, 2011. **36**(6): p. 338-45.
9. Jimenez, J.L., et al., *Cryo-electron microscopy structure of an SH3 amyloid fibril and model of the molecular packing.* EMBO J, 1999. **18**(4): p. 815-21.
10. Toyama, B.H. and J.S. Weissman, *Amyloid structure: conformational diversity and consequences.* Annu Rev Biochem, 2011. **80**: p. 557-85.
11. Volpatti, L.R., et al., *A clear view of polymorphism, twist, and chirality in amyloid fibril formation.* ACS Nano, 2013. **7**(12): p. 10443-8.
12. Fitzpatrick, A.W., et al., *Atomic structure and hierarchical assembly of a cross-beta amyloid fibril.* Proc Natl Acad Sci U S A, 2013. **110**(14): p. 5468-73.
13. Tycko, R., et al., *Evidence for novel beta-sheet structures in Iowa mutant beta-amyloid fibrils.* Biochemistry, 2009. **48**(26): p. 6072-84.
14. Kajava, A.V., U. Baxa, and A.C. Steven, *Beta arcades: recurring motifs in naturally occurring and disease-related amyloid fibrils.* FASEB J, 2010. **24**(5): p. 1311-9.
15. Petkova, A.T., et al., *Self-propagating, molecular-level polymorphism in Alzheimer's beta-amyloid fibrils.* Science, 2005. **307**(5707): p. 262-5.
16. Lu, J.X., et al., *Molecular structure of beta-amyloid fibrils in Alzheimer's disease brain tissue.* Cell, 2013. **154**(6): p. 1257-68.

17. Nilsson, K.P., et al., *Imaging distinct conformational states of amyloid-beta fibrils in Alzheimer's disease using novel luminescent probes.* ACS Chem Biol, 2007. **2**(8): p. 553-60.
18. Nystrom, S., et al., *Evidence for age-dependent in vivo conformational rearrangement within Abeta amyloid deposits.* ACS Chem Biol, 2013. **8**(6): p. 1128-33.
19. Sikorski, P. and E. Atkins, *New model for crystalline polyglutamine assemblies and their connection with amyloid fibrils.* Biomacromolecules, 2005. **6**(1): p. 425-32.
20. Sharma, D., et al., *Polyglutamine homopolymers having 8-45 residues form slablike beta-crystallite assemblies.* Proteins, 2005. **61**(2): p. 398-411.
21. Sivanandam, V.N., et al., *The aggregation-enhancing huntingtin N-terminus is helical in amyloid fibrils.* J Am Chem Soc, 2011. **133**(12): p. 4558-66.
22. Schneider, R., et al., *Structural Characterization of Polyglutamine Fibrils by Solid-State NMR Spectroscopy.* J Mol Biol, 2011.
23. Bugg, C.W., et al., *Structural features and domain organization of huntingtin fibrils.* J Biol Chem, 2012. **287**(38): p. 31739-46.
24. Thakur, A.K. and R. Wetzel, *Mutational analysis of the structural organization of polyglutamine aggregates.* Proc Natl Acad Sci U S A, 2002. **99**(26): p. 17014-9.
25. Kar, K., et al., *beta-hairpin-mediated nucleation of polyglutamine amyloid formation.* J Mol Biol, 2013. **425**(7): p. 1183-97.
26. Perutz, M.F., et al., *Amyloid fibers are water-filled nanotubes.* Proceedings of the National Academy of Sciences, 2002. **99**(8): p. 5591-5595.
27. Parton, R.G., M. Hanzal-Bayer, and J.F. Hancock, *Biogenesis of caveolae: a structural model for caveolin-induced domain formation.* J Cell Sci, 2006. **119**(Pt 5): p. 787-96.
28. Anderson, R.G., *The caveolae membrane system.* Annu Rev Biochem, 1998. **67**: p. 199-225.
29. Parton, R.G. and K. Simons, *The multiple faces of caveolae.* Nat Rev Mol Cell Biol, 2007. **8**(3): p. 185-94.
30. Murata, M., et al., *VIP21/caveolin is a cholesterol-binding protein.* Proc Natl Acad Sci U S A, 1995. **92**(22): p. 10339-43.
31. Schlegel, A., et al., *A role for the caveolin scaffolding domain in mediating the membrane attachment of caveolin-1. The caveolin scaffolding domain is both necessary and sufficient for membrane binding in vitro.* J Biol Chem, 1999. **274**(32): p. 22660-7.
32. Epand, R.M., B.G. Sayer, and R.F. Epand, *Caveolin scaffolding region and cholesterol-rich domains in membranes.* J Mol Biol, 2005. **345**(2): p. 339-50.
33. Couet, J., M. Sargiacomo, and M.P. Lisanti, *Interaction of a receptor tyrosine kinase, EGF-R, with caveolins. Caveolin binding negatively regulates tyrosine and serine/threonine kinase activities.* J Biol Chem, 1997. **272**(48): p. 30429-38.
34. Andrew, E.R., A. Bradbury, and R.G. Eades, *Removal of Dipolar Broadening of Nuclear Magnetic Resonance Spectra of Solids by Specimen Rotation.* Nature, 1959. **183**(4678): p. 1802-1803.
35. Takegoshi, K., S. Nakamura, and T. Terao, *13C–1H dipolar-assisted rotational resonance in magic-angle spinning NMR.* Chemical Physics Letters, 2001. **344**(5–6): p. 631-637.
36. Baldus, M., et al., *Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems.* Molecular Physics, 1998. **95**(6): p. 1197-1207.

37. Wishart, D.S., B.D. Sykes, and F.M. Richards, *The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy*. Biochemistry, 1992. **31**(6): p. 1647-51.
38. Hafsa, N. and D. Wishart, *CSI 2.0: a significantly improved version of the Chemical Shift Index*. Journal of Biomolecular NMR, 2014. **60**(2-3): p. 131-146.
39. Shen, Y., et al., *TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts.* Journal of Biomolecular NMR, 2009. **44**(4): p. 213-223.
40. Veshtort, M. and R.G. Griffin, *SPINEVOLUTION: A powerful tool for the simulation of solid and liquid state NMR experiments.* Journal of Magnetic Resonance, 2006. **178**(2): p. 248-282.
41. Costa, P.R., et al., *Solid-state NMR measurement of Psi in peptides: a NCCN 2Q-heteronuclear local field experiment.* Chemical Physics Letters, 1997. **280**(1-2): p. 95-103.
42. Gullion, T. and J. Schaefer, *Rotational-echo double-resonance NMR.* Journal of Magnetic Resonance (1969), 1989. **81**(1): p. 196-200.
43. Feng, X., et al., *Direct determination of a molecular torsional angle by solid-state NMR.* Chemical Physics Letters, 1996. **257**(3–4): p. 314-320.
44. Ross, C.A., *Polyglutamine Pathogenesis: Emergence of Unifying Mechanisms for Huntington's Disease and Related Disorders.* Neuron, 2002. **35**(5): p. 819-822.
45. Hoop, C.L., et al., *Polyglutamine amyloid core boundaries and flanking domain dynamics in huntingtin fragment fibrils determined by solid-state nuclear magnetic resonance.* Biochemistry, 2014. **53**(42): p. 6653-66.
46. Sahoo, B., et al., *Aggregation behavior of chemically synthesized, full-length huntingtin exon1*. Biochemistry, 2014. **53**(24): p. 3897-907.
47. Rockabrand, E., et al., *The first 17 amino acids of Huntingtin modulate its sub-cellular localization, aggregation and effects on calcium homeostasis.* Hum Mol Genet, 2007. **16**(1): p. 61-77.
48. Tam, S., et al., *The chaperonin TRiC blocks a huntingtin sequence element that promotes the conformational switch to aggregation.* Nat Struct Mol Biol, 2009. **16**(12): p. 1279-85.
49. Thakur, A.K., et al., *Polyglutamine disruption of the huntingtin exon 1 N terminus triggers a complex aggregation mechanism.* Nat Struct Mol Biol, 2009. **16**(4): p. 380-389.
50. Zhang, H., S. Neal, and D.S. Wishart, *RefDB: a database of uniformly referenced protein chemical shifts.* J Biomol NMR, 2003. **25**(3): p. 173-95.
51. Kricheldorf, H.R. and D. Müller, *Secondary structure of peptides: 15. 13C n.m.r. CP/MAS study of solid elastin and proline-containing copolyesters.* International Journal of Biological Macromolecules, 1984. **6**(3): p. 145-151.
52. Liebman, S.W. and S.C. Meredith, *Protein folding: sticky N17 speeds huntingtin pile-up.* Nat Chem Biol, 2010. **6**(1): p. 7-8.
53. Gu, X., et al., *Serines 13 and 16 are critical determinants of full-length human mutant huntingtin induced disease pathogenesis in HD mice.* Neuron, 2009. **64**(6): p. 828-40.
54. Mishra, R., et al., *Serine phosphorylation suppresses huntingtin amyloid accumulation by altering protein aggregation properties.* J Mol Biol, 2012. **424**(1-2): p. 1-14.
55. De Gortari, I., et al., *Time averaging of NMR chemical shifts in the MLF peptide in the solid state.* J Am Chem Soc, 2010. **132**(17): p. 5993-6000.

56. Tycko, R., *Solid-state NMR studies of amyloid fibril structure.* Annu Rev Phys Chem, 2011. **62**: p. 279-99.
57. Zuccato, C., M. Valenza, and E. Cattaneo, *Molecular Mechanisms and Potential Therapeutical Targets in Huntington's Disease.* Physiological Reviews, 2010. **90**(3): p. 905-981.
58. Bates, G.P., Benn, C., *The polyglutmamine diseases*, in *Huntington's Disease*, G.P. Bates, Harper, P.S., Jones, L. , Editor. 2002, Oxford Press: Oxford, UK. p. 429-472.
59. Kar, K., et al., *D-polyglutamine amyloid recruits L-polyglutamine monomers and kills cells.* J Mol Biol, 2014. **426**(4): p. 816-29.
60. Pieri, L., et al., *Fibrillar alpha-synuclein and huntingtin exon 1 assemblies are toxic to the cells.* Biophys J, 2012. **102**(12): p. 2894-905.
61. Yang, W., et al., *Aggregated polyglutamine peptides delivered to nuclei are toxic to mammalian cells.* Hum Mol Genet, 2002. **11**(23): p. 2905-17.
62. Nekooki-Machida, Y., et al., *Distinct conformations of in vitro and in vivo amyloids of huntingtin-exon1 show different cytotoxicity.* Proc Natl Acad Sci U S A, 2009. **106**(24): p. 9679-84.
63. Arrasate, M. and S. Finkbeiner, *Protein aggregates in Huntington's disease.* Exp Neurol, 2012. **238**(1): p. 1-11.
64. Chen, S., et al., *Amyloid-like Features of Polyglutamine Aggregates and Their Assembly Kinetics†.* Biochemistry, 2002. **41**(23): p. 7391-7399.
65. Bhattacharyya, A., et al., *Oligoproline Effects on Polyglutamine Conformation and Aggregation.* Journal of Molecular Biology, 2006. **355**(3): p. 524-535.
66. Darnell, G., et al., *Flanking Polyproline Sequences Inhibit β-Sheet Structure in Polyglutamine Segments by Inducing PPII-like Helix Structure.* Journal of Molecular Biology, 2007. **374**(3): p. 688-704.
67. Darnell, G.D., et al., *Mechanism of Cis-Inhibition of PolyQ Fibrillation by PolyP: PPII Oligomers and the Hydrophobic Effect.* Biophysical Journal, 2009. **97**(8): p. 2295-2305.
68. DiFiglia, M., et al., *Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain.* Science, 1997. **277**(5334): p. 1990-3.
69. Lunkes, A., et al., *Proteases acting on mutant huntingtin generate cleaved products that differentially build up cytoplasmic and nuclear inclusions.* Mol Cell, 2002. **10**(2): p. 259-69.
70. Sathasivam, K., et al., *Aberrant splicing of HTT generates the pathogenic exon 1 protein in Huntington disease.* Proc Natl Acad Sci U S A, 2013. **110**(6): p. 2366-70.
71. Mangiarini, L., et al., *Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice.* Cell, 1996. **87**(3): p. 493-506.
72. Chow, W.N., et al., *Degradation of mutant huntingtin via the ubiquitin/proteasome system is modulated by FE65.* Biochem J, 2012. **443**(3): p. 681-9.
73. Burke, K.A., et al., *The interaction of polyglutamine peptides with lipid membranes is regulated by flanking sequences associated with huntingtin.* J Biol Chem, 2013. **288**(21): p. 14993-5005.
74. Qin, Z.H., et al., *Huntingtin bodies sequester vesicle-associated proteins by a polyproline-dependent interaction.* J Neurosci, 2004. **24**(1): p. 269-81.
75. Duennwald, M.L., et al., *Flanking sequences profoundly alter polyglutamine toxicity in yeast.* Proc Natl Acad Sci U S A, 2006. **103**(29): p. 11045-50.

76. Duennwald, M.L., et al., *A network of protein interactions determines polyglutamine toxicity.* Proc Natl Acad Sci U S A, 2006. **103**(29): p. 11051-6.
77. Hollenbach, B., et al., *Aggregation of truncated GST-HD exon 1 fusion proteins containing normal range and expanded glutamine repeats.* Philos Trans R Soc Lond B Biol Sci, 1999. **354**(1386): p. 991-4.
78. Wetzel and Mishra, *Structural BiologyOrder, Disorder, and Conformational Flux*. 'Oxford University Press': Oxford, UK.
79. Jayaraman, M., et al., *Slow Amyloid Nucleation via α-Helix-Rich Oligomeric Intermediates in Short Polyglutamine-Containing Huntingtin Fragments.* Journal of Molecular Biology, 2012. **415**(5): p. 881-899.
80. Chellgren, B.W. and T.P. Creamer, *Short sequences of non-proline residues can adopt the polyproline II helical conformation.* Biochemistry, 2004. **43**(19): p. 5864-9.
81. Chellgren, B.W., A.F. Miller, and T.P. Creamer, *Evidence for polyproline II helical structure in short polyglutamine tracts.* J Mol Biol, 2006. **361**(2): p. 362-71.
82. Jayaraman, M., et al., *Kinetically competing huntingtin aggregation pathways control amyloid polymorphism and properties.* Biochemistry, 2012. **51**(13): p. 2706-16.
83. Kim, M.W., et al., *Secondary structure of Huntingtin amino-terminal region.* Structure, 2009. **17**(9): p. 1205-12.
84. Dlugosz, M. and J. Trylska, *Secondary structures of native and pathogenic huntingtin N-terminal fragments.* J Phys Chem B, 2011. **115**(40): p. 11597-608.
85. Davies, P., et al., *Consequences of poly-glutamine repeat length for the conformation and folding of the androgen receptor amino-terminal domain.* J Mol Endocrinol, 2008. **41**(5): p. 301-14.
86. Laghaei, R. and N. Mousseau, *Spontaneous formation of polyglutamine nanotubes with molecular dynamics simulations.* J Chem Phys, 2010. **132**(16): p. 165102.
87. Nakano, M., et al., *Comparative characterization of short monomeric polyglutamine peptides by replica exchange molecular dynamics simulation.* J Phys Chem B, 2010. **114**(20): p. 7056-61.
88. Williamson, T.E., et al., *Modulation of polyglutamine conformations and dimer formation by the N-terminus of huntingtin.* J Mol Biol, 2010. **396**(5): p. 1295-309.
89. Tobelmann, M.D. and R.M. Murphy, *Location trumps length: polyglutamine-mediated changes in folding and aggregation of a host protein.* Biophys J, 2011. **100**(11): p. 2773-82.
90. Stanley, C.B., T. Perevozchikova, and V. Berthelier, *Structural formation of huntingtin exon 1 aggregates probed by small-angle neutron scattering.* Biophys J, 2011. **100**(10): p. 2504-12.
91. Jayaraman, M., et al., *Kinetically competing huntingtin aggregation pathways control amyloid polymorphism and properties.* Biochemistry, 2012. **51**(13): p. 2706-2716.
92. Burke, K.A., et al., *The interaction of polyglutamine peptides with lipid membranes is regulated by flanking sequences associated with huntingtin.* Journal of Biological Chemistry, 2013. **288**(21): p. 14993-15005.
93. Sahoo, B., et al., *Aggregation Behavior of Chemically Synthesized, Full-Length Huntingtin Exon1*. Biochemistry, 2014. **53**(24): p. 3897-3907.
94. Perevozchikova, T., et al., *Investigating the Structural Impact of the Glutamine Repeat in Huntingtin Assembly.* Biophysical Journal, 2014. **107**(2): p. 411-421.

95. O'Nuallain, B., et al., *Kinetics and Thermodynamics of Amyloid Assembly Using a High-Performance Liquid Chromatography–Based Sedimentation Assay*, in *Methods in Enzymology*, K. Indu and W. Ronald, Editors. 2006, Academic Press. p. 34-74.
96. Poirier, M.A., et al., *Huntingtin spheroids and protofibrils as precursors in polyglutamine fibrilization.* J Biol Chem, 2002. **277**(43): p. 41032-7.
97. Harris, R.K., et al., *Further conventions for NMR shielding and chemical shifts (IUPAC Recommendations 2008).* Magn Reson Chem, 2008. **46**(6): p. 582-98.
98. Bennett, A.E., et al., *Heteronuclear decoupling in rotating solids.* The Journal of Chemical Physics, 1995. **103**(16): p. 6951-6958.
99. Torchia, D.A. and A. Szabo, *Spin-Lattice Relaxation in Solids.* J Magn Reson, 1982. **49**(1): p. 107-121.
100. Lewandowski, J.R., et al., *Structural complexity of a composite amyloid fibril.* J Am Chem Soc, 2011. **133**(37): p. 14686-14698.
101. Munowitz, M., et al., *Two-dimensional rotational spin-echo nuclear magnetic resonance in solids: correlation of chemical shift and dipolar interactions.* J Am Chem Soc, 1981. **103**(10): p. 2529-2533.
102. Carravetta, M., et al., *Symmetry principles for the design of radiofrequency pulse sequences in the nuclear magnetic resonance of rotating solids.* Chem Phys Lett, 2000. **321**(3–4): p. 205-215.
103. Zhao, X., M. Edén, and M.H. Levitt, *Recoupling of heteronuclear dipolar interactions in solid-state NMR using symmetry-based pulse sequences.* Chemical Physics Letters, 2001. **342**(3–4): p. 353-361.
104. Veshtort, M. and R.G. Griffin, *SPINEVOLUTION: a powerful tool for the simulation of solid and liquid state NMR experiments.* J Magn Reson, 2006. **178**(2): p. 248-282.
105. Kumashiro, K.K., et al., *A Novel Tool for Probing Membrane Protein Structure: Solid-State NMR with Proton Spin Diffusion and X-Nucleus Detection.* Journal of the American Chemical Society, 1998. **120**(20): p. 5043-5051.
106. Andronesi, O.C., et al., *Characterization of Alzheimer's-like Paired Helical Filaments from the Core Domain of Tau Protein Using Solid-State NMR Spectroscopy.* Journal of the American Chemical Society, 2008. **130**(18): p. 5922-5928.
107. Huster, D., X. Yao, and M. Hong, *Membrane Protein Topology Probed by 1H Spin Diffusion from Lipids Using Solid-State NMR Spectroscopy.* Journal of the American Chemical Society, 2002. **124**(5): p. 874-883.
108. Darnell, G., et al., *Flanking polyproline sequences inhibit beta-sheet structure in polyglutamine segments by inducing PPII-like helix structure.* J Mol Biol, 2007. **374**(3): p. 688-704.
109. Darnell, G.D., et al., *Mechanism of cis-inhibition of polyQ fibrillation by polyP: PPII oligomers and the hydrophobic effect.* Biophys J, 2009. **97**(8): p. 2295-2305.
110. Chellgren, B.W., A.-F. Miller, and T.P. Creamer, *Evidence for polyproline II helical structure in short polyglutamine tracts.* J Mol Biol, 2006. **361**(2): p. 362-371.
111. Chellgren, B.W. and T.P. Creamer, *Short Sequences of Non-Proline Residues Can Adopt the Polyproline II Helical Conformation.* Biochemistry, 2004. **43**(19): p. 5864-5869.
112. Candel, A.M., et al., *The high-resolution NMR structure of a single-chain chimeric protein mimicking a SH3–peptide complex.* FEBS Letters, 2007. **581**(4): p. 687-692.

113. Kricheldorf, H.R. and D. Müller, *Secondary structure of peptides: 15. 13C nmr CP/MAS study of solid elastin and proline-containing copolyesters.* Int J Biol Macromol, 1984. **6**(3): p. 145-151.
114. Kar, K., et al., *β-hairpin-mediated nucleation of polyglutamine amyloid formation.* J Mol Biol, 2013. **425**(7): p. 1183-1197.
115. Schneider, R., et al., *Structural Characterization of Polyglutamine Fibrils by Solid-State NMR Spectroscopy.* J Mol Biol, 2011. **412**(1): p. 121-136.
116. Qin, Z.-H., et al., *Huntingtin bodies sequester vesicle-associated proteins by a polyproline-dependent interaction.* J Neurosci, 2004. **24**(1): p. 269-281.
117. Southwell, A.L., et al., *Intrabodies Binding the Proline-Rich Domains of Mutant Huntingtin Increase Its Turnover and Reduce Neurotoxicity.* J Neurosci, 2008. **28**(36): p. 9013-9020.
118. Chow, W.N.V., et al., *Degradation of mutant huntingtin via the ubiquitin/proteasome system is modulated by FE65.* Biochem J, 2012. **443**(3): p. 681-689.
119. Gao, Y.-G., et al., *Autoinhibitory Structure of the WW Domain of HYPB/SETD2 Regulates Its Interaction with the Proline-Rich Region of Huntingtin.* Structure, 2014. **22**(3): p. 378-386.
120. Ader, C., et al., *Structural rearrangements of membrane proteins probed by water-edited solid-state NMR spectroscopy.* J Am Chem Soc, 2009. **131**(1): p. 170-176.
121. Shen, Y., et al., *TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts.* J Biomol NMR, 2009. **44**(4): p. 213-223.
122. Lewandowski, J.R., et al., *Site-Specific Measurement of Slow Motions in Proteins.* Journal of the American Chemical Society, 2011. **133**(42): p. 16762-16765.
123. Kalk, A. and H.J.C. Berendsen, *Proton magnetic relaxation and spin diffusion in proteins.* Journal of Magnetic Resonance (1969), 1976. **24**(3): p. 343-366.
124. Chevelkov, V., et al., *Combined analysis of (15)N relaxation data from solid- and solution-state NMR spectroscopy.* J Am Chem Soc, 2007. **129**(42): p. 12594-12595.
125. Fry, E.A., et al., *CSA-Enabled Spin Diffusion Leads to MAS Rate-Dependent T_1's at High Field.* J Am Chem Soc, 2011. **133**(5): p. 1156-1158.
126. Helmus, J.J., et al., *Conformational flexibility of Y145Stop human prion protein amyloid fibrils probed by solid-state nuclear magnetic resonance spectroscopy.* J Am Chem Soc, 2010. **132**(7): p. 2393-2403.
127. Hou, G., et al., *Multidimensional magic angle spinning NMR spectroscopy for site-resolved measurement of proton chemical shift anisotropy in biological solids.* J Am Chem Soc, 2013. **135**(4): p. 1358-1368.
128. Yao, L., et al., *The Impact of Hydrogen Bonding on Amide 1H Chemical Shift Anisotropy Studied by Cross-Correlated Relaxation and Liquid Crystal NMR Spectroscopy.* J Am Chem Soc, 2010. **132**(31): p. 10866-10875.
129. Kelly, M.A., et al., *Host−Guest Study of Left-Handed Polyproline II Helix Formation.* Biochemistry, 2001. **40**(48): p. 14376-14383.
130. Michalek, M., E.S. Salnikov, and B. Bechinger, *Structure and Topology of the Huntingtin 1-17 Membrane Anchor by a Combined Solution and Solid-State NMR Approach.* Biophys J, 2013. **105**(3): p. 699-710.
131. Kim, M.W., et al., *Secondary structure of Huntingtin amino-terminal region.* Structure, 2009. **17**(9): p. 1205-1212.

132. Atwal, R.S., et al., *Huntingtin has a membrane association signal that can modulate huntingtin aggregation, nuclear entry and toxicity.* Hum Mol Genet, 2007. **16**(21): p. 2600-2615.
133. Williamson, T.E., et al., *Modulation of polyglutamine conformations and dimer formation by the N-terminus of huntingtin.* Journal of molecular biology, 2010. **396**(5): p. 1295-1309.
134. Długosz, M. and J. Trylska, *Secondary structures of native and pathogenic huntingtin N-terminal fragments.* The journal of physical chemistry B, 2011. **115**(40): p. 11597-11608.
135. Jayaraman, M., et al., *Slow amyloid nucleation via alpha-helix-rich oligomeric intermediates in short polyglutamine-containing huntingtin fragments.* J Mol Biol, 2012. **415**(5): p. 881-899.
136. Nagai, Y., et al., *A toxic monomeric conformer of the polyglutamine protein.* Nat Struct Mol Biol, 2007. **14**(4): p. 332-340.
137. Popiel, H.A., et al., *Disruption of the toxic conformation of the expanded polyglutamine stretch leads to suppression of aggregate formation and cytotoxicity.* Biochem Biophys Res Commun, 2004. **317**(4): p. 1200-1206.
138. Margittai, M. and R. Langen, *Fibrils with parallel in-register structure constitute a major class of amyloid fibrils: molecular insights from electron paramagnetic resonance spectroscopy.* Quarterly reviews of biophysics, 2008. **41**(3-4): p. 265-297.
139. van der Wel, P.C.A., *Domain swapping and amyloid fibril conformation.* Prion, 2012. **6**(3): p. 211-216.
140. Tycko, R. and R.B. Wickner, *Molecular Structures of Amyloid and Prion Fibrils: Consensus versus Controversy.* Acc Chem Res, 2013. **46**(7): p. 1487–1496.
141. Wasmer, C., et al., *Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core.* Science (New York, NY), 2008. **319**(5869): p. 1523-1526.
142. Qiang, W., et al., *Antiparallel β-sheet architecture in Iowa-mutant β-amyloid fibrils.* Proceedings Of The National Academy Of Sciences Of The United States Of America, 2012. **109**(12): p. 4443-4448.
143. Kopito, R.R. and D. Ron, *Conformational disease.* Nature cell biology, 2000. **2**(11): p. E207-9.
144. Ferrigno, P. and P.A. Silver, *Polyglutamine expansions: Proteolysis, chaperones, and the dangers of promiscuity.* Neuron, 2000. **26**(1): p. 9-12.
145. Slepko, N., et al., *Normal-repeat-length polyglutamine peptides accelerate aggregation nucleation and cytotoxicity of expanded polyglutamine proteins.* Proceedings Of The National Academy Of Sciences Of The United States Of America, 2006. **103**(39): p. 14367-14372.
146. Wishart, D.S., et al., *1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects.* J Biomol NMR, 1995. **5**(1): p. 67-81.
147. Ko, J., S. Ou, and P.H. Patterson, *New anti-huntingtin monoclonal antibodies: implications for huntingtin conformation and its binding proteins.* Brain Res Bull, 2001. **56**(3-4): p. 319-329.
148. Khoshnan, A., J. Ko, and P.H. Patterson, *Effects of intracellular expression of anti-huntingtin antibodies of various specificities on mutant huntingtin aggregation and toxicity.* Proc Natl Acad Sci USA, 2002. **99**(2): p. 1002-1007.

149. Legleiter, J., et al., *Monoclonal antibodies recognize distinct conformational epitopes formed by polyglutamine in a mutant huntingtin fragment.* J Biol Chem, 2009. **284**(32): p. 21647-21658.
150. Safar, J., et al., *scrapie Amyloid (Prion) Protein Has the Conformational Characteristics of an Aggregated Molten Globule Folding Intermediate.* Biochemistry, 1994. **33**(27): p. 8375-8383.
151. Staniforth, R.A., et al., *Three-dimensional domain swapping in the folded and molten-globule states of cystatins, an amyloid-forming structural superfamily.* The EMBO Journal, 2001. **20**(17): p. 4774-4781.
152. Skora, L., S. Becker, and M. Zweckstetter, *Molten Globule Precursor States Are Conformationally Correlated to Amyloid Fibrils of Human β-2-Microglobulin.* Journal of the American Chemical Society, 2010. **132**(27): p. 9223-9225.
153. Bemporad, F. and F. Chiti, *Protein Misfolded Oligomers: Experimental Approaches, Mechanism of Formation, and Structure-Toxicity Relationships.* Chem Biol, 2012. **19**(3): p. 315-327.
154. Gerber, R., et al., *Oligomerization of the human prion protein proceeds via a molten globule intermediate.* Journal of Biological Chemistry, 2006. **282**(9): p. 6300-6307.
155. Narayanan, S. and B. Reif, *Characterization of chemical exchange between soluble and aggregated states of beta-amyloid by solution-state NMR upon variation of salt conditions.* Biochemistry, 2005. **44**(5): p. 1444-1452.
156. Vendrely, C., et al., *Assembly of the full-length recombinant mouse prion protein I. Formation of soluble oligomers.* Biochimica et biophysica acta, 2005. **1724**(3): p. 355-366.
157. Comellas, G. and C.M. Rienstra, *Protein Structure Determination by Magic-Angle Spinning Solid-State NMR, and Insights into the Formation, Structure, and Stability of Amyloid Fibrils.* Annu Rev Biophys, 2012. **42**(1): p. 515-536.
158. van der Wel, P.C., J.R. Lewandowski, and R.G. Griffin, *Structural characterization of GNNQQNY amyloid fibrils by magic angle spinning NMR.* Biochemistry, 2010. **49**(44): p. 9457-69.
159. Nucifora, L.G., et al., *Identification of novel potentially toxic oligomers formed in vitro from mammalian-derived expanded huntingtin exon-1 protein.* J Biol Chem, 2012. **287**(19): p. 16017-28.
160. Buchanan, L.E., et al., *Structural motif of polyglutamine amyloid fibrils discerned with mixed-isotope infrared spectroscopy.* Proc Natl Acad Sci U S A, 2014. **111**(16): p. 5796-801.
161. Perutz, M.F., et al., *Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases.* Proc Natl Acad Sci U S A, 1994. **91**(12): p. 5355-8.
162. Sen, S., et al., *Role of histidine interruption in mitigating the pathological effects of long polyglutamine stretches in SCA1: A molecular approach.* Protein Sci, 2003. **12**(5): p. 953-962.
163. Thakur, A.K., W. Yang, and R. Wetzel, *Inhibition of polyglutamine aggregate cytotoxicity by a structure-based elongation inhibitor.* FASEB J, 2004. **18**(7): p. 923-925.
164. Bhattacharyya, A.M., A.K. Thakur, and R. Wetzel, *Polyglutamine aggregation nucleation: thermodynamics of a highly unfavorable protein folding reaction.* Proc Natl Acad Sci USA, 2005. **102**(43): p. 15400-15405.

165. Poirier, M.A., H. Jiang, and C.A. Ross, *A structure-based analysis of huntingtin mutant polyglutamine aggregation and toxicity: evidence for a compact beta-sheet structure.* Hum Mol Genet, 2005. **14**(6): p. 765-774.
166. Zhang, Q.C., et al., *A Compact Beta Model of huntingtin Toxicity.* Journal of Biological Chemistry, 2011. **286**(10): p. 8188-8196.
167. Stanger, H.E. and S.H. Gellman, *Rules for Antiparallel β-Sheet Design: d-Pro-Gly Is Superior to l-Asn-Gly for β-Hairpin Nucleation1*. Journal of the American Chemical Society, 1998. **120**(17): p. 4236-4237.
168. Daggett, V., *Alpha-sheet: The toxic conformer in amyloid diseases?* Acc Chem Res, 2006. **39**(9): p. 594-602.
169. Bertini, I., et al., *On the use of ultracentrifugal devices for sedimented solute NMR.* J Biomol NMR, 2012. **54**(2): p. 123-7.
170. Gardiennet, C., et al., *A sedimented sample of a 59 kDa dodecameric helicase yields high-resolution solid-state NMR spectra.* Angew Chem Int Ed Engl, 2012. **51**(31): p. 7855-8.
171. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes.* J Biomol NMR, 1995. **6**(3): p. 277-93.
172. Jaroniec, C.P., et al., *Molecular conformation of a peptide fragment of transthyretin in an amyloid fibril.* Proc Natl Acad Sci U S A, 2002. **99**(26): p. 16748-53.
173. Hohwy, M., C.M. Rienstra, and R.G. Griffin, *Band-selective homonuclear dipolar recoupling in rotating solids.* J Chem Phys, 2002. **117**: p. 4973-4987.
174. Hohwy, M., et al., *Fivefold symmetric homonuclear dipolar recoupling in rotating solids: Application to double quantum spectroscopy.* The Journal of Chemical Physics, 1999. **110**(16): p. 7983-7992.
175. Bajaj, V.S., P.C. van der Wel, and R.G. Griffin, *Observation of a low-temperature, dynamically driven structural transition in a polypeptide by solid-state NMR spectroscopy.* J Am Chem Soc, 2009. **131**(1): p. 118-28.
176. Edwards, R., et al., *Measurement of multiple torsional angles from one-dimensional solid-state NMR spectra: application to the conformational analysis of a ligand in its biological receptor site.* Phys Chem Chem Phys, 2010. **12**(42): p. 13999-4008.
177. Lewandowski, J.R., et al., *Structural Complexity of a Composite Amyloid Fibril.* Journal of the American Chemical Society, 2011. **133**(37): p. 14686-14698.
178. Lewandowski, J.R., et al., *Anisotropic collective motion contributes to nuclear spin relaxation in crystalline proteins.* J Am Chem Soc, 2010. **132**(4): p. 1246-1248.
179. Babin, V., C. Roland, and C. Sagui, *The alpha-sheet: a missing-in-action secondary structure?* Proteins, 2011. **79**(3): p. 937-46.
180. Schubert, M., et al., *A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on 13C chemical shift statistics.* J Biomol NMR, 2002. **24**(2): p. 149-54.
181. Richardson, J.S., *The Anatomy and Taxonomy of Protein Structure*, in *Advances in Protein Chemistry*, J.T.E. C.B. Anfinsen and M.R. Frederic, Editors. 1981, Academic Press. p. 167-339.
182. Armen, R.S., D.O. Alonso, and V. Daggett, *Anatomy of an amyloidogenic intermediate: conversion of beta-sheet to alpha-sheet structure in transthyretin at acidic pH.* Structure, 2004. **12**(10): p. 1847-63.

183. Armen, R.S., et al., *Pauling and Corey's alpha-pleated sheet structure may define the prefibrillar amyloidogenic intermediate in amyloid disease.* Proc Natl Acad Sci U S A, 2004. **101**(32): p. 11622-7.
184. Milner-White, J.E., et al., *Amyloid formation may involve alpha- to beta sheet interconversion via peptide plane flipping.* Structure, 2006. **14**(9): p. 1369-76.
185. Sambashivan, S., et al., *Amyloid-like fibrils of ribonuclease A with three-dimensional domain-swapped and native-like structure.* Nature, 2005. **437**(7056): p. 266-9.
186. Jaroniec, C.P., et al., *High-resolution molecular structure of a peptide in an amyloid fibril determined by magic angle spinning NMR spectroscopy.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(3): p. 711-716.
187. Hong, M., et al., *Structure of an elastin-mimetic polypeptide by solid-state NMR chemical shift analysis.* Biopolymers, 2003. **70**(2): p. 158-168.
188. Yao, X.L. and M. Hong, *Structure Distribution in an Elastin-Mimetic Peptide (VPGVG)3 Investigated by Solid-State NMR.* J Am Chem Soc, 2004. **126**(13): p. 4199-4210.
189. Venkatachalam, C.M., *Stereochemical criteria for polypeptides and proteins. V. Conformation of a system of three linked peptide units.* Biopolymers, 1968. **6**(10): p. 1425-1436.
190. Chou, P.Y. and G.D. Fasman, *β-turns in proteins.* Journal of Molecular Biology, 1977. **115**(2): p. 135-175.
191. Pal, D. and P. Chakrabarti, *Cis peptide bonds in proteins: residues involved, their conformations, interactions and locations.* J Mol Biol, 1999. **294**(1): p. 271-288.
192. Gunasekaran, K., et al., *Conformational interconversions in peptide β-turns: analysis of turns in proteins and computational estimates of barriers.* J Mol Biol, 1998. **284**(5): p. 1505-1516.
193. Karle, I.L., S.K. Awasthi, and P. Balaram, *A designed beta-hairpin peptide in crystals.* Proc Natl Acad Sci USA, 1996. **93**(16): p. 8189-8193.
194. Stanger, H.E., et al., *Length-dependent stability and strand length limits in antiparallel -sheet secondary structure.* Proc. Natl. Acad. Sci. U.S.A., 2001. **98**(21): p. 12015-12020.
195. Reinert, Z.E., G.A. Lengyel, and W.S. Horne, *Protein-like Tertiary Folding Behavior from Heterogeneous Backbones.* J Am Chem Soc, 2013. **135**(34): p. 12528-12531.
196. Haque, T.S., J.C. Little, and S.H. Gellman, *Stereochemical requirements for β-hairpin formation: model studies with four-residue peptides and depsipeptides.* J Am Chem Soc, 1996. **118**(29): p. 6975-6985.
197. Razani, B., S.E. Woodman, and M.P. Lisanti, *Caveolae: from cell biology to animal physiology.* Pharmacol Rev, 2002. **54**(3): p. 431-67.
198. Anderson, R.G.W., *The caveolae membrane system.* Annu Rev Biochem, 1998. **67**: p. 199-225.
199. Engelman, J.A., et al., *Molecular genetics of the caveolin gene family: implications for human cancers, diabetes, Alzheimer disease, and muscular dystrophy.* Am J Hum Genet, 1998. **63**(6): p. 1578-87.
200. Ostermeyer, A.G., et al., *Role of the hydrophobic domain in targeting caveolin-1 to lipid droplets.* J Cell Biol, 2004. **164**(1): p. 69-78.
201. Aoki, S., et al., *The role of proline in the membrane re-entrant helix of caveolin-1.* J Biol Chem, 2010. **285**(43): p. 33371-80.

202. Li, S., K.S. Song, and M.P. Lisanti, *Expression and characterization of recombinant caveolin. Purification by polyhistidine tagging and cholesterol-dependent incorporation into defined lipid membranes.* J Biol Chem, 1996. **271**(1): p. 568-73.
203. Dietzen, D.J., W.R. Hastings, and D.M. Lublin, *Caveolin is palmitoylated on multiple cysteine residues. Palmitoylation is not necessary for localization of caveolin to caveolae.* J Biol Chem, 1995. **270**(12): p. 6838-42.
204. Uittenbogaard, A. and E.J. Smart, *Palmitoylation of caveolin-1 is required for cholesterol binding, chaperone complex formation, and rapid transport of cholesterol to caveolae.* The Journal of biological chemistry, 2000. **275**(33): p. 25595-9.
205. Li, S., J. Couet, and M.P. Lisanti, *Src tyrosine kinases, Galpha subunits, and H-Ras share a common membrane-anchored scaffolding protein, caveolin. Caveolin binding negatively regulates the auto-activation of Src tyrosine kinases.* J Biol Chem, 1996. **271**(46): p. 29182-90.
206. Couet, J., et al., *Identification of peptide and protein ligands for the caveolin-scaffolding domain. Implications for the interaction of caveolin with caveolae-associated proteins.* J Biol Chem, 1997. **272**(10): p. 6525-33.
207. Murata, M., et al., *VIP21/caveolin is a cholesterol-binding protein.* Proc Natl Acad Sci USA, 1995. **92**(22): p. 10339-43.
208. Le Lan, C., J.M. Neumann, and N. Jamin, *Role of the membrane interface on the conformation of the caveolin scaffolding domain: a CD and NMR study.* FEBS Lett, 2006. **580**(22): p. 5301-5.
209. Levin, A.M., et al., *Exploring the interaction between the protein kinase A catalytic subunit and caveolin-1 scaffolding domain with shotgun scanning, oligomer complementation, NMR, and docking.* Protein Sci, 2006. **15**(3): p. 478-86.
210. Fernandez, I., et al., *Mechanism of caveolin filament assembly.* Proc Natl Acad Sci USA, 2002. **99**(17): p. 11193-8.
211. Le Lan, C., et al., *Structural and dynamic properties of juxta-membrane segments of caveolin-1 and caveolin-2 at the membrane interface.* Eur Biophys J, 2010. **39**(2): p. 307-25.
212. Levin, A.M., et al., *Double barrel shotgun scanning of the caveolin-1 scaffolding domain.* ACS Chem Biol, 2007. **2**(7): p. 493-500.
213. Jamin, N., et al., *Characterization of the cholesterol recognition amino acid consensus sequence of the peripheral-type benzodiazepine receptor.* Mol Endocrinol, 2005. **19**(3): p. 588-94.
214. Ortegren, U., et al., *Lipids and glycosphingolipids in caveolae and surrounding plasma membrane of primary rat adipocytes.* Eur J Biochem, 2004. **271**(10): p. 2028-36.
215. Spisni, E., et al., *Structural insights into the function of human caveolin 1.* Biochem Biophys Res Commun, 2005. **338**(3): p. 1383-90.
216. McDermott, A., *Structure and dynamics of membrane proteins by magic angle spinning solid-state NMR.* Annu. Rev. Biophys., 2009. **38**: p. 385-403.
217. Hong, M., *Oligomeric structure, dynamics, and orientation of membrane proteins from solid-state NMR.* Structure, 2006. **14**(12): p. 1731-40.
218. Opella, S.J. and F.M. Marassi, *Structure determination of membrane proteins by NMR spectroscopy.* Chem Rev, 2004. **104**(8): p. 3587-606.
219. Smith, S.O., et al., *Implications of threonine hydrogen bonding in the glycophorin A transmembrane helix dimer.* Biophys J, 2002. **82**(5): p. 2476-86.

220. Bennett, A.E., et al., *Heteronuclear decoupling in rotating solids.* J Chem Phys, 1995. **103**(16): p. 6951-6958.
221. Luo, W. and M. Hong, *Conformational changes of an ion channel detected through water-protein interactions using solid-state NMR spectroscopy.* J Am Chem Soc, 2010. **132**(7): p. 2378-84.
222. Forbes, J., et al., *Some new developments in solid-state nuclear magnetic resonance spectroscopic studies of lipids and biological membranes, including the effects of cholesterol in model and natural systems.* J. Chem. Soc., Faraday Trans. 1, 1988. **84**: p. 3821-3849.
223. Guo, W. and J.A. Hamilton, *13C MAS NMR studies of crystalline cholesterol and lipid mixtures modeling atherosclerotic plaques.* Biophys J, 1996. **71**(5): p. 2857-68.
224. Goddard, T.G. and D.G. Kneller, *SPARKY 3*. University of California, San Fransisco
225. Vranken, W.F., et al., *The CCPN data model for NMR spectroscopy: development of a software pipeline.* Proteins, 2005. **59**(4): p. 687-96.
226. Harris, R.K., et al., *Further conventions for NMR shielding and chemical shifts (IUPAC Recommendations 2008).* Magn Reson Chem, 2008. **46**(6): p. 582-98.
227. Jones, D.T., *Protein secondary structure prediction based on position-specific scoring matrices.* J Mol Biol, 1999. **292**(2): p. 195-202.
228. Rost, B., G. Yachdav, and J. Liu, *The PredictProtein server.* Nucleic Acids Res, 2004. **32**(Web Server issue): p. W321-6.
229. Montgomerie, S., et al., *PROTEUS2: a web server for comprehensive protein structure prediction and structure-based annotation.* Nucleic Acids Res, 2008. **36**(Web Server issue): p. W202-9.
230. Dufourc, E.J., et al., *Dynamics of phosphate head groups in biomembranes. Comprehensive analysis using phosphorus-31 nuclear magnetic resonance lineshape and relaxation time measurements.* Biophys J, 1992. **61**(1): p. 42-57.
231. Van der Wel, P.C.A., et al., *Tryptophan-anchored transmembrane peptides promote formation of nonlamellar phases in phosphatidylethanolamine model membranes in a mismatch-dependent manner.* Biochemistry, 2000. **39**(11): p. 3124-3133.
232. Costello, A.L. and T.M. Alam, *Investigating the impact of cholesterol on magnetically aligned sphingomyelin/cholesterol multilamellar vesicles using static (31)P NMR.* Chem Phys Lipids, 2010. **163**(6): p. 506-13.
233. Holland, G.P., S.K. McIntyre, and T.M. Alam, *Distinguishing individual lipid headgroup mobility and phase transitions in raft-forming lipid mixtures with 31P MAS NMR.* Biophys J, 2006. **90**(11): p. 4248-60.
234. Huang, J. and G.W. Feigenson, *A microscopic interaction model of maximum solubility of cholesterol in lipid bilayers.* Biophys J, 1999. **76**(4): p. 2142-57.
235. Benetis, N., et al., *Static CP P-31 NMR multilamellar bilayer broadlines in the absence and presence of the bioactive dipeptide beta-Ala-Tyr or Glu.* Chem Phys, 2005. **314**(1-3): p. 57-72.
236. Seelig, J., *P-31 nuclear magnetic-resonance and head group structure of phospholipids in membranes.* Biochim Biophys Acta, 1978. **515**(2): p. 105-140.
237. Kong, J. and S. Yu, *Fourier transform infrared spectroscopic analysis of protein secondary structures.* Acta Biochim Biophys Sin (Shanghai), 2007. **39**(8): p. 549-59.

238. Reis, O., R. Winter, and T.W. Zerda, *The effect of high external pressure on DPPC-cholesterol multilamellar vesicles: a pressure-tuning Fourier transform infrared spectroscopy study.* Biochim. Biophys. Acta, 1996. **1279**: p. 5-16.
239. Wishart, D.S. and B.D. Sykes, *The 13C Chemical-Shift Index: A simple method for the identification of protein secondary structure using 13C chemical-shift data* J. Biomol. NMR, 1994. **4**: p. 171-180.
240. Zhang, H., S. Neal, and D.S. Wishart, *RefDB: a database of uniformly referenced protein chemical shifts.* J. Biomol. NMR, 2003. **25**(3): p. 173-95.
241. Senes, A., D.E. Engel, and W.F. DeGrado, *Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs.* Curr Opin Struct Biol, 2004. **14**(4): p. 465-79.
242. Morris, R., et al., *Rafts, little caves and large potholes: how lipid structure interacts with membrane proteins to create functionally diverse membrane environments.* Subcell Biochem, 2004. **37**: p. 35-118.
243. Volonte, D. and F. Galbiati, *Inhibition of thioredoxin reductase 1 by caveolin 1 promotes stress-induced premature senescence.* EMBO Rep, 2009. **10**(12): p. 1334-40.
244. Scolari, S., et al., *Interaction of mammalian seminal plasma protein PDC-109 with cholesterol - Implications for a putative CRAC domain.* Biochemistry, 2010. **49**: p. 9027-9031.
245. Wah, D.A., et al., *Sperm coating mechanism from the 1.8 A crystal structure of PDC-109-phosphorylcholine complex.* Structure, 2002. **10**(4): p. 505-14.
246. Lewandowski, J.R., et al., *Structural Complexity of a Composite Amyloid Fibril.* J Am Chem Soc, 2011. **133**(37): p. 14686-98.
247. Li, J., et al., *Amyloid-like fibrils from a domain-swapping protein feature a parallel, in-register conformation without native-like interactions.* J Biol Chem, 2011. **286**(33): p. 28988-28995.
248. Li, H. and V. Papadopoulos, *Peripheral-type benzodiazepine receptor function in cholesterol transport. Identification of a putative cholesterol recognition/interaction amino acid sequence and consensus pattern.* Endocrinology, 1998. **139**(12): p. 4991-7.
249. Epand, R.M., *Proteins and cholesterol-rich domains.* Biochim Biophys Acta, 2008. **1778**(7-8): p. 1576-82.
250. Coleman, D.E. and S.R. Sprang, *Crystal structures of the G protein Gi alpha 1 complexed with GDP and Mg2+: a crystallographic titration experiment.* Biochemistry, 1998. **37**(41): p. 14376-85.
251. Li, S., et al., *Evidence for a regulated interaction between heterotrimeric G proteins and caveolin.* J Biol Chem, 1995. **270**(26): p. 15693-701.